# Supplementary Appendix for
# ‘Tail-robust factor modelling of vector and tensor time series in high dimensions’

Matteo Barigozzi[1] and Haeran Cho[2] and Hyeyoung Maeng[3]

December 6, 2025

## Contents



[1]Department of Economics, Università di Bologna. Email: matteo.barigozzi@unibo.it.
[2]School of Mathematics, University of Bristol. Email: haeran.cho@bristol.ac.uk.
[3]Department of Mathematical Sciences, Durham University. Email: hyeyoung.maeng@durham.ac.uk.

## A Asymptotic identifiability

From (2), we have the mode-$k$ second moment matrix $\mathbf{\Gamma}^{(k)}$ decomposed as

$$\begin{aligned}\mathbf{\Gamma}^{(k)} &= \frac{1}{p_{-k}}\mathbf{\Lambda}_k \left(\frac{1}{n}\sum_{t=1}^{n} \mathsf{mat}_k(\mathcal{F}_t)\mathbf{\Delta}_k^\top \mathbf{\Delta}_k \mathsf{mat}_k(\mathcal{F}_t)^\top\right) \mathbf{\Lambda}_k^\top + \frac{1}{np_{-k}}\sum_{t=1}^{n} \mathsf{E}\left[\mathsf{mat}_k(\boldsymbol{\xi}_t)\mathsf{mat}_k(\boldsymbol{\xi}_t)^\top\right] \\ &= \mathbf{\Lambda}_k \left(\frac{1}{n}\sum_{t=1}^{n} \mathsf{mat}_k(\mathcal{F}_t)\mathsf{mat}_k(\mathcal{F}_t)^\top\right) \mathbf{\Lambda}_k^\top + \frac{1}{np_{-k}}\sum_{t=1}^{n} \mathsf{E}\left[\mathsf{mat}_k(\boldsymbol{\xi}_t)\mathsf{mat}_k(\boldsymbol{\xi}_t)^\top\right] \\ &=: \mathbf{\Gamma}_\chi^{(k)} + \mathbf{\Gamma}_\xi^{(k)}. \end{aligned} \tag{A.1}$$

Let us define $\widetilde{\mathbf{\Gamma}}_\chi^{(k)} := \mathbf{\Lambda}_k \mathbf{\Gamma}_f^{(k)} \mathbf{\Lambda}_k^\top$. Under Assumption 1 (i), we have for all $k \in [K]$,

$$\frac{1}{p_{-k}}\mathbf{\Delta}_k^\top \mathbf{\Delta}_k = \mathbf{I}_{r_{-k}}$$

for all $p_k$'s. This, together with Assumption 2, leads to

$$\frac{1}{p_k}\left\|\mathbf{\Gamma}_\chi^{(k)} - \widetilde{\mathbf{\Gamma}}_\chi^{(k)}\right\| \le \frac{1}{p_k}\left\|\mathbf{\Lambda}_k \left(\frac{1}{n}\sum_{t=1}^{n} \mathsf{mat}_k(\mathcal{F}_t)\mathsf{mat}_k(\mathcal{F}_t)^\top - \mathbf{\Gamma}_f^{(k)}\right) \mathbf{\Lambda}_k^\top\right\| = o(1). \tag{A.2}$$

Assumptions 1 (i) and 2 indicate that the $r_k$ non-zero eigenvalues of $\widetilde{\mathbf{\Gamma}}_\chi^{(k)}$ are distinct and diverging linearly in $p_k$ as $p_k \to \infty$, which is inherited by the $r_k$ leading eigenvalues of $\mathbf{\Gamma}_\chi^{(k)}$ from (A.2) and Weyl's inequality (see Lemma A.1 below). From this, it follows that the $r_k$ common factors are pervasive across the $p_k$ cross-sections of $\mathsf{mat}_k(\boldsymbol{\chi}_t)$ for all $k \in [K]$. On the other hand, $\mathbf{\Gamma}_\xi^{(k)}$ has bounded eigenvalues for all $p_k$, either under Assumptions 4 or 5, see Lemma A.2. Then, thanks to Weyl's inequality, these observations lead to a diverging gap between the leading $r_k$ eigenvalues of $\mathbf{\Gamma}^{(k)}$ and the remainder, which ensures that the latent components $\boldsymbol{\chi}_t$ and $\boldsymbol{\xi}_t$ are separable in all $K$ modes asymptotically as $p_k \to \infty$.
For any non-negative definite matrix $\mathbf{A}$, let $\mu_j(\mathbf{A})$ denote its $j$-th largest eigenvalue. Also, let us write $\mu_{f,j}^{(k)} = \mu_j(\mathbf{\Gamma}_f^{(k)})$.

**Lemma A.1.** *Suppose that Assumptions 1–2 hold. Then for each $k \in [K]$, the $r_k$ non-zero eigenvalues of $\mathbf{\Gamma}_\chi^{(k)}$ and $\widetilde{\mathbf{\Gamma}}_\chi^{(k)}$, denoted by $\mu_{\chi,j}^{(k)}$ and $\widetilde{\mu}_{\chi,j}^{(k)}$, $j \in [r_k]$, respectively, satisfy*

$$\begin{aligned}\left|\frac{1}{p_k}\widetilde{\mu}_{\chi,j}^{(k)} - \mu_{f,j}^{(k)}\right| &= o(1), \\ \left|\frac{1}{p_k}\mu_{\chi,j}^{(k)} - \mu_{f,j}^{(k)}\right| &= o(1).\end{aligned}$$

*Proof.* By Assumption 1 (i),

$$\frac{1}{p_k}\mathbf{\Gamma}_f^{(k)}\mathbf{\Lambda}_k^\top \mathbf{\Lambda}_k = \mathbf{\Gamma}_f^{(k)}$$

and thus for all $j \in [r_k]$,

$$\frac{1}{p_k}\mu_j\left(\mathbf{\Gamma}_f^{(k)}\mathbf{\Lambda}_k^\top\mathbf{\Lambda}_k\right) = \mu_{f,j}^{(k)}.$$

Since $r_k$ non-zero eigenvalues of $\widetilde{\mathbf{\Gamma}}_\chi^{(k)}$ are identical to those of $\mathbf{\Gamma}_f^{(k)}\mathbf{\Lambda}_k^\top\mathbf{\Lambda}_k$, the first statement holds. The second statement follows from evoking (A.2) with the first one. □

**Lemma A.2.** *Suppose that Assumption 3 (i) and either of Assumptions 4 (i) or 5 (i) hold. Then, there exists a constant $C_\epsilon > 0$ that may depend on $\epsilon$, such that $\|\mathbf{\Gamma}_\xi^{(k)}\| \le C_\epsilon\omega^2$ for all $k \in [K]$.*

*Proof.* Firstly, under Assumptions 4 (i), $\mathbf{\Gamma}_\xi^{(k)}$ is diagonal such that

$$\left\|\mathbf{\Gamma}_\xi^{(k)}\right\| \le \max_{i\in[p_k]}\frac{1}{np_{-k}}\sum_{t=1}^n\sum_{\mathbf{i}\in\prod_{k'\in[K]}[p_{k'}]:\, i_k=i}\|\xi_{\mathbf{i},t}\|_2^2 \le \omega^2.$$

Next, let us write the elements of $\mathsf{mat}_k(\boldsymbol{\xi}_t)$ by $\xi_{k,i\ell,t}$ for $i \in [p_k]$ and $\ell \in [p_{-k}]$, and denote by $\|\cdot\|_1$ the matrix norm induced by the vector $\ell_1$-norm. Under Assumption 5 (i), we have for all $i \in [p_k]$,

$$\begin{aligned}\|\mathbf{\Gamma}_\xi^{(k)}\|_1 &\le \max_{i\in[p_k]}\frac{1}{np_{-k}}\sum_{j\in[p_k]}\sum_{t\in[n]}\sum_{\ell\in[p_{-k}]}\left|\mathsf{E}\left(\xi_{k,i\ell,t}\xi_{k,j\ell,t}\right)\right| \\ &\le \max_{i\in[p_k]}\frac{8}{np_{-k}}\sum_{j\in[p_k]}\sum_{t\in[n]}\sum_{\ell\in[p_{-k}]}\|\xi_{k,i\ell,t}\|_{2+2\epsilon}\|\xi_{k,j\ell,t}\|_{2+2\epsilon}\exp\left(-\frac{c_0\epsilon|i-j|}{1+\epsilon}\right) \\ &\le \max_{i\in[p_k]}8\omega^2\sum_{j\in[p_k]}\exp\left(-\frac{c_0\epsilon|i-j|}{1+\epsilon}\right) \le C_\epsilon\omega^2,\end{aligned}$$

where the first inequality follows from Lemma B.3 and the second from Assumption 3 (i). This establishes that $\|\mathbf{\Gamma}_\xi^{(k)}\| \le \|\mathbf{\Gamma}_\xi^{(k)}\|_1 \le C_\epsilon\omega^2$. □

# B Proofs

Throughout, for a matrix $\mathbf{A} = [a_{ii'}] \in \mathbb{R}^{m\times n}$, we write its Frobenius norm by $\|\mathbf{A}\|_F = |\mathbf{A}|_2$ and write $|\mathbf{A}|_\infty = \max_{i\in[m]}\max_{i'\in[n]}|a_{ii'}|$. By $\mathbf{O}$ and $\mathbf{0}$, we denote a matrix or a vector of zeros whose dimensions depend on the context. We write $a_n \lesssim b_n$ and $a_n = O(b_n)$ interchangeably. Also, with some abuse of notation, we write $\mathbf{A} = \mathbf{B} + o(1)$ for two matrices $\mathbf{A}$ and $\mathbf{B}$ of compatible, fixed dimensions, if the equality holds element-wise.
We write the pairs of eigenvalues and eigenvectors of $\mathbf{\Gamma}^{(k)}$ by $(\mu_j^{(k)}, \mathbf{e}_j^{(k)})$, with $\mu_1^{(k)} \ge \ldots \ge \mu_{p_k}^{(k)}$, and similarly define $(\mu_{\chi,j}^{(k)}, \mathbf{e}_{\chi,j}^{(k)})$ for $\mathbf{\Gamma}_\chi^{(k)}$ in (A.1). We also write $\mathbf{E}_k = [\mathbf{e}_1^{(k)}, \ldots, \mathbf{e}_{r_k}^{(k)}]$ and $\mathbf{E}_{\chi,k} = [\mathbf{e}_{\chi,1}^{(k)}, \ldots, \mathbf{e}_{\chi,r_k}^{(k)}]$. [!!!] Assumption 2 indicates that there exist pairs of fixed, positive constants

$(\alpha_j^{(k)}, \beta_j^{(k)})$, $j \in [r_k]$, such that $\beta_1^{(k)} \geq \mu_{f,1}^{(k)} \geq \alpha_1^{(k)} > \beta_2^{(k)} \geq \ldots \geq \alpha_{r_k-1}^{(k)} > \beta_{r_k}^{(k)} \geq \mu_{f,r_k}^{(k)} \geq \alpha_{r_k}^{(k)}$, where $\mu_{f,j}^{(k)}$ denotes the $j$-th largest eigenvalue of $\mathbf{\Gamma}_f^{(k)}$. We frequently write

$$\mathbf{X}_{k,t} := \mathsf{mat}_k(\mathcal{X}_t) = [X_{k,ii',t},\, i \in [p_k],\, i' \in [p_{-k}]], \quad \text{and}$$
$$\mathbf{X}_{k,t}^{\mathsf{t}}(\tau) := \mathsf{mat}_k(\mathcal{X}_t^{\mathsf{t}}(\tau)) = [X_{k,ii',t}^{\mathsf{t}}(\tau),\, i \in [p_k],\, i' \in [p_{-k}]],$$

and $\mathbf{Z}_{k,t}(\tau) = \mathbf{X}_{k,t}^{\mathsf{t}}(\tau) - \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}}(\tau))$. By $C_\epsilon > 0$, we denote a constant that depends only on $\epsilon$ which may differ from one instance to another. We write $r_{-k} = r/r_k$ with $r = \prod_{k=1}^K r_k$.

## B.1 Preliminary lemmas

**Lemma B.1.** *Suppose that Assumptions 1 (ii) and 3 hold. Then for any* $\mathbf{i}, \mathbf{i}' \in \prod_{k=1}^K [p_k]$ *and* $t \in [n]$*:*

*(i)* $\|X_{\mathbf{i},t}^{\mathsf{t}}(\tau)\|_\nu \leq \|X_{\mathbf{i},t}\|_\nu \lesssim |\mathcal{F}_t|_2 + \omega$ *for any* $\nu \in [1, 2+2\epsilon]$.

*(ii)* $\mathsf{E}(|X_{\mathbf{i},t}^{\mathsf{t}}(\tau) X_{\mathbf{i}',t}^{\mathsf{t}}(\tau)|^\nu) \lesssim \tau^{2(\nu-1-\epsilon)}(|\mathcal{F}_t|_2^{2+2\epsilon} + \omega^{2+2\epsilon})$ *for any* $\nu \in [2, \infty)$.

*(iii)* $|\mathsf{E}(X_{\mathbf{i},t}^{\mathsf{t}}(\tau)) - \mathsf{E}(X_{\mathbf{i},t})| \lesssim \tau^{-1-2\epsilon}(|\mathcal{F}_t|_2^{2+2\epsilon} + \omega^{2+2\epsilon})$.

*(iv)* $|\mathsf{E}(X_{\mathbf{i},t}^{\mathsf{t}}(\tau) X_{\mathbf{i}',t}^{\mathsf{t}}(\tau)) - \mathsf{E}(X_{\mathbf{i},t} X_{\mathbf{i}',t})| \lesssim \tau^{-2\epsilon}(|\mathcal{F}_t|_2^{2+2\epsilon} + \omega^{2+2\epsilon})$.

*Proof.* For (i), the first inequality follows by construction. Under Assumption 1 (ii),

$$|\chi_{i_1\ldots i_K,t}| = \left| \sum_{j_1 \in [r_1]} \cdots \sum_{j_K \in [r_K]} f_{j_1\ldots j_K,t} \cdot \prod_{k=1}^K \lambda_{i_k j_k} \right| \leq \prod_{k=1}^K r_k \bar{\lambda}^K |\mathcal{F}_t|_2 \tag{B.1}$$

by Cauchy-Schwarz inequality, for any $(i_1, \ldots, i_K)^\top \in \prod_{k=1}^K [p_k]$. Combined with Assumption 3 (i), we have $\|X_{\mathbf{i},t}\|_\nu \leq |\chi_{\mathbf{i},t}| + \|\xi_{\mathbf{i},t}\|_\nu \lesssim |\mathcal{F}_t|_2 + \omega$ by Minkowski inequality.
For (ii), note that

$$\begin{aligned}
\mathsf{E}\left(\left|X_{\mathbf{i},t}^{\mathsf{t}}(\tau) X_{\mathbf{i}',t}^{\mathsf{t}}(\tau)\right|^\nu\right) &\leq \mathsf{E}\left[\min\left(\left|X_{\mathbf{i},t}^{\mathsf{t}}(\tau) X_{\mathbf{i}',t}^{\mathsf{t}}(\tau)\right|^\nu, \tau^{2\nu}\right)\right] \\
&= \tau^{2\nu} \mathsf{E}\left[\min\left(\left|\frac{X_{\mathbf{i},t}^{\mathsf{t}}(\tau) X_{\mathbf{i}',t}^{\mathsf{t}}(\tau)}{\tau^2}\right|^\nu, 1\right)\right] \\
&\leq \tau^{2\nu} \mathsf{E}\left(\left|\frac{X_{\mathbf{i},t}^{\mathsf{t}}(\tau) X_{\mathbf{i}',t}^{\mathsf{t}}(\tau)}{\tau^2}\right|^{1+\epsilon}\right) \leq \tau^{2(\nu-1-\epsilon)} \sqrt{\|X_{\mathbf{i},t}\|_{2+2\epsilon}^{2+2\epsilon} \|X_{\mathbf{i}',t}\|_{2+2\epsilon}^{2+2\epsilon}} \\
&\leq C_\epsilon \tau^{2(\nu-1-\epsilon)}(|\mathcal{F}_t|_2^{2+2\epsilon} + \omega^{2+2\epsilon}),
\end{aligned}$$

where the last inequality follows from (i) and $C_r$ inequality.
For (iii), we have

$$\left|\mathsf{E}(X_{\mathbf{i},t}^{\mathsf{t}}(\tau)) - \mathsf{E}(X_{\mathbf{i},t})\right| \leq \mathsf{E}\left[(X_{\mathbf{i},t} - \mathsf{sign}(X_{\mathbf{i},t})) \cdot \mathbb{I}_{\{|X_{\mathbf{i},t}| > \tau\}}\right] \leq \mathsf{E}\left[|X_{\mathbf{i},t}| \cdot \mathbb{I}_{\{|X_{\mathbf{i},t}| > \tau\}}\right]$$

$$\leq \|X_{\mathbf{i},t}\|_{2+2\epsilon} \cdot \mathsf{P}(|X_{\mathbf{i},t}| > \tau)^{\frac{1+2\epsilon}{2+2\epsilon}} \leq \tau^{-1-2\epsilon}\|X_{\mathbf{i},t}\|_{2+2\epsilon}^{2+2\epsilon}$$
$$\leq C_\epsilon \tau^{-1-2\epsilon}(|\mathcal{F}_t|_2^{2+2\epsilon} + \omega^{2+2\epsilon}),$$

which follows from Hölder's, Markov and $C_r$ inequalities combined with (i).

For (iv), note that for all $\mathbf{i}, \mathbf{i}' \in \prod_{k=1}^K [p_k]$ and $t \in [n]$, we have

$$\left|X_{\mathbf{i},t}^{\mathsf{t}} X_{\mathbf{i}',t}^{\mathsf{t}} - X_{\mathbf{i},t}X_{\mathbf{i}',t}\right| \leq \left|\left(X_{\mathbf{i},t}X_{\mathbf{i}',t} - \mathsf{sign}(X_{\mathbf{i},t}X_{\mathbf{i}',t})\tau^2\right)\mathbb{I}_{\{|X_{\mathbf{i},t}|>\tau\}}\mathbb{I}_{\{|X_{\mathbf{i}',t}|>\tau\}}\right|$$
$$+ \left|X_{\mathbf{i},t}\mathbb{I}_{\{|X_{\mathbf{i},t}|\leq\tau\}}\left(X_{\mathbf{i}',t} - \mathsf{sign}(X_{\mathbf{i}',t})\tau\right)\mathbb{I}_{\{|X_{\mathbf{i}',t}|>\tau\}}\right| + \left|X_{\mathbf{i}',t}\mathbb{I}_{\{|X_{\mathbf{i}',t}|\leq\tau\}}\left(X_{\mathbf{i},t} - \mathsf{sign}(X_{\mathbf{i},t})\tau\right)\mathbb{I}_{\{|X_{\mathbf{i},t}|>\tau\}}\right|$$
$$\leq |X_{\mathbf{i},t}X_{\mathbf{i}',t}|\left(\mathbb{I}_{\{|X_{\mathbf{i},t}|>\tau\}}\mathbb{I}_{\{|X_{\mathbf{i}',t}|>\tau\}} + \mathbb{I}_{\{|X_{\mathbf{i},t}|>\tau\}}\mathbb{I}_{\{|X_{\mathbf{i}',t}|\leq\tau\}} + \mathbb{I}_{\{|X_{\mathbf{i},t}|\leq\tau\}}\mathbb{I}_{\{|X_{\mathbf{i}',t}|>\tau\}}\right)$$
$$\leq |X_{\mathbf{i},t}X_{\mathbf{i}',t}|\left(\mathbb{I}_{\{|X_{\mathbf{i},t}|>\tau\}} + \mathbb{I}_{\{|X_{\mathbf{i}',t}|>\tau\}}\right).$$

Then by (i), we have

$$\left|\mathsf{E}(X_{\mathbf{i},t}^{\mathsf{t}}(\tau)X_{\mathbf{i}',t}^{\mathsf{t}}(\tau)) - \mathsf{E}(X_{\mathbf{i},t}X_{\mathbf{i}',t})\right| \leq \mathsf{E}\left[\left|X_{\mathbf{i},t}X_{\mathbf{i}',t}\right|\left(\mathbb{I}_{\{|X_{\mathbf{i},t}|>\tau\}} + \mathbb{I}_{\{|X_{\mathbf{i}',t}|>\tau\}}\right)\right]$$
$$\leq \|X_{\mathbf{i},t}\|_{2+2\epsilon}\|X_{\mathbf{i}',t}\|_{2+2\epsilon}\left[\mathsf{P}(|X_{\mathbf{i},t}| > \tau) + \mathsf{P}(|X_{\mathbf{i}',t}| > \tau)\right]^{\frac{\epsilon}{1+\epsilon}}$$
$$\lesssim \|X_{\mathbf{i},t}\|_{2+2\epsilon}\|X_{\mathbf{i}',t}\|_{2+2\epsilon} \cdot \frac{\|X_{\mathbf{i},t}\|_{2+2\epsilon}^{2\epsilon} + \|X_{\mathbf{i}',t}\|_{2+2\epsilon}^{2\epsilon}}{\tau^{2\epsilon}}$$
$$\leq C_\epsilon \tau^{-2\epsilon}(|\mathcal{F}_t|_2^{2+2\epsilon} + \omega^{2+2\epsilon}).$$

□

**Lemma B.2** (Lemma 1 of Wang and Tsay (2023)). *For any $\alpha$-mixing time series $\{Y_t\}_{t\in\mathbb{Z}}$ and measurable function $f(\cdot)$, the sequence of the transformed process $\{f(Y_t)\}_{t\in\mathbb{Z}}$ is also $\alpha$-mixing with its mixing coefficients bounded by those of the original sequence.*

**Lemma B.3** (Corollary A.2 of Hall and Heyde (1980)). *Suppose that $X$ and $Y$ are random variables which are $\mathcal{G}$ and $\mathcal{H}$-measurable, respectively, for $\sigma$-algebra $\mathcal{G}$ and $\mathcal{H}$, and that $\|X\|_{\nu_1}, \|Y\|_{\nu_2} < \infty$ for some $\nu_1, \nu_2 \in (1,\infty)$ with $\nu_1^{-1} + \nu_2^{-1} < 1$. Then,*

$$|\mathsf{Cov}(X,Y)| \leq 8\|X\|_{\nu_1}\|Y\|_{\nu_2}[\alpha(\mathcal{G},\mathcal{H})]^{1-\nu_1^{-1}-\nu_2^{-1}},$$

*where $\alpha(\mathcal{G},\mathcal{H}) = \sup_{A\in\mathcal{G},B\in\mathcal{H}} |\mathsf{P}(A\cap B) - \mathsf{P}(A)\mathsf{P}(B)|$.*

**Lemma B.4.** *Let $\mathbf{S}, \widetilde{\mathbf{S}}, \widehat{\mathbf{S}} \in \mathbb{R}^{p\times p}$ denote symmetric, non-negative definite matrices fulfilling (a subset of) the following conditions.*

*(C1) $\mathbf{S}$ has $r$ non-zero eigenvalues $\mu_j$, $j \in [r]$ with $r \leq p$, satisfying $|p^{-1}\mu_j - \gamma_j| = o(1)$, where*

$$\beta_1 \geq \frac{\gamma_1}{p} \geq \alpha_1 > \beta_2 \geq \ldots \geq \alpha_{r-1} > \beta_r \geq \frac{\gamma_r}{p} \geq \alpha_r > 0,$$

*with some pairs of positive constants $(\alpha_j, \beta_j)$, $j \in [r]$.*

*(C2)* $p^{-1}\|\widehat{\mathbf{S}} - \widetilde{\mathbf{S}}\|_F = O_P(\zeta_{n,p})$ *where* $\zeta_{n,p} \to 0$ *as* $n, p \to \infty$.

*(C3)* $\|\widetilde{\mathbf{S}} - \mathbf{S}\| \leq C_1 < \infty$.

*(C4)* $\max(|\mathbf{S}|_\infty, |\widetilde{\mathbf{S}}|_\infty) \leq C_2 < \infty$.

*Denoting the pairs of eigenvalues and eigenvectors of* $\widehat{\mathbf{S}}$ *by* $(\widehat{\mu}_j, \widehat{\mathbf{e}}_j)$, $j \geq 1$, *let us write* $\widehat{\mathbf{M}} = diag(\widehat{\mu}_j,\, j \in [r])$ *and* $\widehat{\mathbf{E}} = [\widehat{e}_{ij},\, i \in [p],\, j \in [r]]$, *and analogously define* $(\mu_j, \mathbf{e}_j)$, $\mathbf{M}$ *and* $\mathbf{E}$ *(resp.* $(\widetilde{\mu}_j, \widetilde{\mathbf{e}}_j)$, $\widetilde{\mathbf{M}}$ *and* $\widetilde{\mathbf{E}}$*) for* $\mathbf{S}$ *(resp.* $\widetilde{\mathbf{S}}$*).*

*(i) Suppose that (C1), (C2) and (C3) hold. Then, there exist diagonal matrices* $\mathbf{J}, \widetilde{\mathbf{J}} \in \mathbb{R}^{p\times p}$ *with* $\pm 1$ *on their diagonal such that*

$$\|\widehat{\mathbf{E}} - \widetilde{\mathbf{E}}\widetilde{\mathbf{J}}\|_F = O_P(\zeta_{n,p}) \quad \textit{and} \quad \|\widehat{\mathbf{E}} - \mathbf{E}\mathbf{J}\|_F = O_P(\zeta_{n,p} \vee p^{-1}).$$

*(ii) Suppose that (C1), (C2) and (C3) hold. Then,*

$$p^{-1}\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\| = O_P(\zeta_{n,p}), \quad p^{-1}\|\widehat{\mathbf{M}} - \mathbf{M}\| = O_P(\zeta_{n,p} \vee p^{-1}),$$
$$p\|\widehat{\mathbf{M}}^{-1} - \widetilde{\mathbf{M}}^{-1}\| = O_P(\zeta_{n,p}) \quad \textit{and} \quad p\|\widehat{\mathbf{M}}^{-1} - \mathbf{M}^{-1}\| = O_P(\zeta_{n,p} \vee p^{-1}).$$

*(iii) Suppose that (C1), (C3) and (C4) hold. Then, there exists a constant* $C_3 > 0$ *such that*

$$\max_{j\in[r]}\max_{i\in[p]} |\widetilde{e}_{ij}| \leq C_3 p^{-1/2} \quad \textit{and} \quad \max_{j\in[r]}\max_{i\in[p]} |e_{ij}| \leq C_3 p^{-1/2}.$$

*Proof.* For (i), note that by Chebyshev's inequality and (C2), we have

$$\frac{1}{p}\|\widehat{\mathbf{S}} - \widetilde{\mathbf{S}}\| \leq \frac{1}{p}\|\widehat{\mathbf{S}} - \widetilde{\mathbf{S}}\|_F = O_P\left(\zeta_{n,p}\right).$$

Also by Weyl's inequality and (C3), we have $|\widetilde{\mu}_j - \mu_j| \leq C_1$ for all $j \in [p]$. Then by Theorem 2 of Yu et al. (2015) and (C1), there exists such $\widetilde{\mathbf{J}}$ satisfying

$$\begin{aligned}\|\widehat{\mathbf{E}} - \widetilde{\mathbf{E}}\widetilde{\mathbf{J}}\|_F &\leq \frac{2\sqrt{2r}\|\widehat{\mathbf{S}} - \widetilde{\mathbf{S}}\|}{\min(\widetilde{\mu}_1 - \widetilde{\mu}_2, \widetilde{\mu}_r - \widetilde{\mu}_{r+1})} \\ &= O_P\left(\frac{p\zeta_{n,p}}{\min((\alpha_1 - \beta_2)p - 2C_1, \alpha_r p - 2C_1)}\right) = O_P(\zeta_{n,p}).\end{aligned}$$

The second result follows from the analogous arguments.
For (ii), by Weyl's inequality and (C2), for all $j \in [p]$,

$$\frac{1}{p}|\widehat{\mu}_j - \widetilde{\mu}_j| \leq \frac{1}{p}\|\widehat{\mathbf{S}} - \widetilde{\mathbf{S}}\| = O_P\left(\zeta_{n,p}\right),$$

from which the first statement follows. Also from (C1) and (C3), $p^{-1}\mu_r \geq \alpha_r + o(1)$ and thus $p^{-1}\widetilde{\mu}_r \geq \alpha_r + o(1) - C_1 p^{-1}$ and $p^{-1}\widehat{\mu}_r \geq \alpha_r + o(1) - C_1 p^{-1} + O_P(\zeta_{n,p})$, which imply that the matrices $p^{-1}\widetilde{\mathbf{M}}$ and $p^{-1}\widehat{\mathbf{M}}$ are asymptotically invertible with

$$\left\| \left(p^{-1}\widetilde{\mathbf{M}}\right)^{-1} \right\| \leq \frac{1}{\alpha_r(1+o(1))} \quad \text{and} \quad \left\| \left(p^{-1}\widehat{\mathbf{M}}\right)^{-1} \right\| \leq \frac{1}{\alpha_r(1+o_P(1))}.$$

Further, we have

$$\left\| \left(p^{-1}\widehat{\mathbf{M}}\right)^{-1} - \left(p^{-1}\widetilde{\mathbf{M}}\right)^{-1} \right\| \leq \left\| \left(p^{-1}\widehat{\mathbf{M}}\right)^{-1} - \left(p^{-1}\widetilde{\mathbf{M}}\right)^{-1} \right\|_F = \sqrt{p^2 \sum_{j\in[r]} \left(\frac{1}{\widehat{\mu}_j} - \frac{1}{\widetilde{\mu}_j}\right)^2}$$
$$\leq \sum_{j\in[r]} \frac{p^{-1}|\widetilde{\mu}_j - \mu_j|}{p^{-1}\widetilde{\mu}_j \cdot p^{-1}\mu_j} = O_P\left(\frac{rp^2}{\mu_r^2}\zeta_{n,p}\right).$$

The second and the fourth claims follow similarly.
To see that (iii) holds, note that

$$\widetilde{s}_{ii} = \sum_{j\in[r]} \widetilde{\mu}_j |\widetilde{e}_{ij}|^2 \leq C_2,$$

which implies that $\max_{j\in[r]} \max_{i\in[p]} |\widetilde{e}_{ij}| \leq C_2/\sqrt{p(\alpha_r + o(1)) - C_1} \leq C_3/\sqrt{p}$ for some $C_3 > 0$. Similarly, the second claim follows.

□

## B.2 Proof of Theorem 1

We first derive an initial rate of estimation for $\widehat{\mathbf{\Gamma}}^{(k)}(\tau)$, from which that of $\widehat{\mathbf{\Lambda}}_k(\tau)$ follows; under Assumption 5, the second claim in (B.3) gives the rate in Theorem 1 together with Lemma B.9.

**Proposition B.1.** *Suppose that Assumptions 1, 2 and 3 hold, as well as either of Assumption 4 or Assumption 5. For each $k \in [K]$, recalling the definition of $\tau_{n,p}^{(k)}$ in* (10)*, we set $\tau \asymp \tau_{n,p}^{(k)}$. Then, there exist diagonal matrices $\mathbf{J}_k, \widehat{\mathbf{J}}_k \in \mathbb{R}^{r_k \times r_k}$ with $\pm 1$ on their diagonal entries such that as* $\min(n, p_1, \ldots, p_K) \to \infty$,

$$\frac{1}{p_k}\left\|\widehat{\mathbf{\Gamma}}^{(k)}(\tau) - \mathbf{\Gamma}_\chi^{(k)}\right\| = O_P\left(\psi_{n,p}^{(k)} \vee \frac{1}{p_k}\right), \tag{B.2}$$

$$\left\|\widehat{\mathbf{E}}_k(\tau) - \mathbf{E}_k\mathbf{J}_k\right\| = O_P\left(\psi_{n,p}^{(k)}\right) \quad \text{and} \quad \left\|\widehat{\mathbf{E}}_k(\tau) - \mathbf{E}_{\chi,k}\widehat{\mathbf{J}}_k\right\| = O_P\left(\psi_{n,p}^{(k)} \vee \frac{1}{p_k}\right), \tag{B.3}$$

*with $\psi_{n,p}^{(k)}$ defined in* (11).

From here on, we make Assumption 4. For each $k \in [K]$, denote by $\widehat{\mathbf{M}}_k(\tau) \in \mathbb{R}^{r_k \times r_k}$ the diagonal matrix containing the eigenvalues $\widehat{\mu}_j^{(k)}(\tau)$, $j \in [r_k]$, of $\widehat{\mathbf{\Gamma}}^{(k)}(\tau)$ on its diagonal. From now on, we suppress the dependence on $\tau$ where there is no confusion. By Weyl's inequality, (B.2), Lemma A.1 and Assumption 2, we have

$$\Lambda_{\max}\left(p_k^{-1}\widehat{\mathbf{M}}_k\right) \le \beta_1^{(k)} + o_P(1) \quad \text{and} \quad \Lambda_{\min}\left(p_k^{-1}\widehat{\mathbf{M}}_k\right) \ge \alpha_{r_k}^{(k)} + o_P(1), \tag{B.4}$$

which ensures the asymptotic invertibility of $p_k^{-1}\widehat{\mathbf{M}}_k$. Also Lemma B.5 shows that

$$\left\| \left(p_k^{-1}\widehat{\mathbf{M}}_k\right)^{-1} \right\| \le \left\| \left(p_k^{-1}\mathbf{M}_{\chi,k}\right)^{-1} \right\| + O_P\left(\psi_{n,p}^{(k)} \vee \frac{1}{p_k}\right) = \frac{1}{\alpha_{r_k}^{(k)}} + o_P(1) = O_P(1), \tag{B.5}$$

where $\mathbf{M}_{\chi,k} = \mathrm{diag}(\mu_{\chi,1}^{(k)}, \ldots, \mu_{\chi,r_k}^{(k)})$. Next, we decompose $\widehat{\mathbf{\Gamma}}^{(k)}$ as

$$\begin{aligned}
\frac{1}{p_k}\widehat{\mathbf{\Gamma}}^{(k)} &= \frac{1}{np}\sum_{t\in[n]} \mathbf{Z}_{k,t}\mathbf{Z}_{k,t}^\top \\
&+ \frac{1}{np}\sum_{t\in[n]} \mathbf{Z}_{k,t}\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top + \frac{1}{np}\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)\mathbf{Z}_{k,t}^\top \\
&+ \frac{1}{np}\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t}\right)^\top + \frac{1}{np}\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t}\right)\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top \\
&+ \frac{1}{np}\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t})\mathsf{E}(\mathbf{X}_{k,t})^\top =: T_1 + T_{2,1} + T_{2,2} + T_{3,1} + T_{3,2} + \frac{1}{p_k}\mathbf{\Gamma}_\chi^{(k)},
\end{aligned} \tag{B.6}$$

where the last equality follows from Assumption 1 (i). Then, noting that $\widehat{\mathbf{\Gamma}}^{(k)}\widehat{\mathbf{E}}_k = \widehat{\mathbf{E}}_k\widehat{\mathbf{M}}_k$, making use of the decomposition in (B.6), we have

$$\begin{aligned}
&\widehat{\mathbf{E}}_k - \frac{1}{\sqrt{p_k}}\mathbf{\Lambda}_k\widehat{\mathbf{H}}_k = (T_1 + T_{2,1} + T_{2,2} + T_{3,1} + T_{3,2})\,\widehat{\mathbf{E}}_k\left(\frac{1}{p_k}\widehat{\mathbf{M}}_k\right)^{-1}, \\
&\text{with} \quad \widehat{\mathbf{H}}_k = \frac{1}{n\sqrt{p_k}}\sum_{t\in[n]} \mathsf{mat}_k(\mathcal{F}_t)\mathsf{mat}_k(\mathcal{F}_t)^\top\mathbf{\Lambda}_k^\top\widehat{\mathbf{E}}_k\left(\frac{1}{p_k}\widehat{\mathbf{M}}_k\right)^{-1},
\end{aligned}$$

by Assumptions 1 (i). This, together with Lemmas B.6–B.8 and (B.5), shows that

$$\frac{1}{p_k}\left\|\widehat{\mathbf{\Gamma}}^{(k)} - \mathbf{\Gamma}_\chi^{(k)}\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p_k} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}}\right), \tag{B.7}$$

$$\left\|\widehat{\mathbf{E}}_k - \frac{1}{\sqrt{p_k}}\mathbf{\Lambda}_k\widehat{\mathbf{H}}_k\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p_k} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}}\right). \tag{B.8}$$

We conclude the proof by noting that

$$\|\widehat{\mathbf{H}}_k\| \le \frac{1}{n}\sum_{t\in[n]} |\mathcal{F}_t|_2^2 \cdot \frac{1}{\sqrt{p_k}}\|\mathbf{\Lambda}_k\|\|\widehat{\mathbf{E}}_k\| \left\|\left(\frac{1}{p_k}\widehat{\mathbf{M}}_k\right)^{-1}\right\| \le (\alpha_{r_k}^{(k)})^{-1}\omega^2(1+o_P(1))$$

by Assumptions 3 (ii) and (B.5), and hence

$$\mathbf{I}_{r_k} = \widehat{\mathbf{E}}_k^\top\widehat{\mathbf{E}}_k = \frac{1}{\sqrt{p_k}}\widehat{\mathbf{E}}_k^\top\mathbf{\Lambda}_k\widehat{\mathbf{H}}_k + o_P(1) = \frac{1}{p_k}\widehat{\mathbf{H}}_k^\top\mathbf{\Lambda}_k^\top\mathbf{\Lambda}_k\widehat{\mathbf{H}}_k + o_P(1) = \widehat{\mathbf{H}}_k^\top\widehat{\mathbf{H}}_k + o_P(1), \quad \text{(B.9)}$$

by Assumptions 1 (i).

#### B.2.1 Proof of Proposition B.1

*Proof of* (B.2). We suppress the dependence on $\tau$ where there is no confusion. Denote by $\mathbf{\Gamma}^{(k)} = [\gamma_{ij}^{(k)}]_{i,j\in[p_k]}$ and $\widehat{\mathbf{\Gamma}}^{(k)} = [\widehat{\gamma}_{ij}^{(k)}]_{i,j\in[p_k]}$. Then,

$$\left|\widehat{\gamma}_{ij}^{(k)} - \gamma_{ij}^{(k)}\right| \le \left|\mathsf{E}(\widehat{\gamma}_{ij}^{(k)}) - \gamma_{ij}^{(k)}\right| + \left|\widehat{\gamma}_{ij}^{(k)} - \mathsf{E}(\widehat{\gamma}_{ij}^{(k)})\right| =: T_{1,ij} + T_{2,ij}.$$

By Lemma B.1 (iv) and Assumption 3 (ii), we have

$$\begin{aligned} T_{1,ij} &\le \frac{1}{np_{-k}}\sum_{t\in[n]}\sum_{i'\in[p_{-k}]} \left|\mathsf{E}(X^{\mathsf{t}}_{k,ii',t}X^{\mathsf{t}}_{k,ji',t}) - \mathsf{E}(X_{k,ii',t}X_{k,ji',t})\right| \\ &\lesssim \frac{1}{n\tau^{2\epsilon}}\sum_{t\in[n]}(|\mathcal{F}_t|_2^{2+2\epsilon} + \omega^{2+2\epsilon}) \le \frac{2\omega^{2+2\epsilon}}{\tau^{2\epsilon}}. \end{aligned} \quad \text{(B.10)}$$

Next, we proceed to bound $T_{2,ij}$ separately under Assumptions 4 and 5.

**Under Assumption 4.** By Lemma B.2, the sequence $\{Y_{ij,\ell,t},\ t\in[n],\ \ell\in[p_{-k}]\}$ formed by concatenating $\{Y_{ij,\ell,t}\}_{t\in[n]}$ with $Y_{ij,\ell,t} = X^{\mathsf{t}}_{k,i\ell,t}X^{\mathsf{t}}_{k,j\ell,t} - \mathsf{E}(X^{\mathsf{t}}_{k,i\ell,t}X^{\mathsf{t}}_{k,j\ell,t})$, is also $\alpha$-mixing with the mixing coefficient as in Assumption 4 (i). Then for some $\nu > 2$, by Lemma B.3, we have

$$|\mathsf{Cov}(Y_{ij,\ell,t}, Y_{ij,\ell,u})| \le 8\|Y_{ij,\ell,t}\|_\nu\|Y_{ij,\ell,u}\|_\nu(\alpha(|t-u|))^{\frac{\nu-2}{\nu}}, \quad \text{where} \quad \text{(B.11)}$$

$$\|Y_{ij,\ell,t}\|_\nu^\nu = \mathsf{E}\left(\left|X^{\mathsf{t}}_{i\ell,t}X^{\mathsf{t}}_{j\ell,t} - \mathsf{E}(X^{\mathsf{t}}_{i\ell,t}X^{\mathsf{t}}_{j\ell,t})\right|^\nu\right) \le 2^\nu\mathsf{E}\left(\left|X^{\mathsf{t}}_{i\ell,t}X^{\mathsf{t}}_{i'j\ell,t}\right|^\nu\right) \quad \text{(B.12)}$$

by $C_r$ inequality. Noting that

$$T_{2,ij} = \frac{1}{np_{-k}}\sum_{\ell\in[p_{-k}]}\sum_{t\in[n]} Y_{ij,\ell,t},$$

we combine (B.11)–(B.12) with Lemma B.1 (ii) and Assumption 4 and obtain

$$\begin{aligned}\|T_{2,ij}\|_2^2 &\le \frac{1}{(np_{-k})^2}\sum_{t,u\in[n]}\sum_{\ell\in[p_{-k}]}|\mathsf{Cov}(Y_{ij,\ell,t},Y_{ij,\ell,u})|\\ &\lesssim \frac{\tau^{\frac{4(\nu-1-\epsilon)}{\nu}}}{n^2p_{-k}}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^{\frac{2+2\epsilon}{\nu}}(|\mathcal{F}_u|_2+\omega)^{\frac{2+2\epsilon}{\nu}}\exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right)\\ &\lesssim \frac{\tau^{2-2\epsilon}}{n^2p_{-k}}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^{1+\epsilon}(|\mathcal{F}_u|_2+\omega)^{1+\epsilon}\exp\left(-\frac{c_0|t-u|}{3\log(np_{-k})}\right)\\ &\lesssim \frac{\omega^{2+2\epsilon}c_\epsilon\tau^{2-2\epsilon}\log(np_{-k})}{np_{-k}},\end{aligned}$$

by setting $\nu = 2+\log^{-1}(np_{-k})$. In third inequality above, we use that

$$\tau^{\frac{4(\nu-1-\epsilon)}{\nu}} \le \tau^{2(1-\epsilon+\log^{-1}(np_{-k}))} \lesssim \tau^{2-2\epsilon}.$$

Noting that the upper bounds on $|T_{1,ij}|$ and $\|T_{2,ij}\|_2$ do not depend on $i$, $j$ or $k$, it follows that

$$\begin{aligned}\frac{1}{p_k}\left\|\widehat{\mathbf{\Gamma}}^{(k)}-\mathbf{\Gamma}^{(k)}\right\|_F &\le \sqrt{\frac{1}{p_k^2}\sum_{i,j\in[p_k]}T_{1,ij}^2}+\frac{1}{p_k}\left\|\widehat{\mathbf{\Gamma}}^{(k)}-\mathsf{E}(\widehat{\mathbf{\Gamma}}^{(k)})\right\|_F\\ &= O_P\left(\frac{\omega^{2+2\epsilon}}{\tau^{2\epsilon}}+\omega^{1+\epsilon}\tau^{1-\epsilon}\sqrt{\frac{\log(np_{-k})}{np_{-k}}}\right)=O_P\left(\omega^2\left(\frac{\log(np_{-k})}{np_{-k}}\right)^{\frac{\epsilon}{1+\epsilon}}\right)\end{aligned}$$

by Markov's inequality, with $\tau \asymp \tau_{n,p}^{(k)}$ in (10).

**Under Assumption 5.** WLOG, for notational convenience, we fix $k=1$ and denote by $\mathbf{i}_{2:K}=(i_2,\ldots,i_K)$. Let us define $Y_{ij,\mathbf{i}_{2:K},t}=X_{1,(i,\mathbf{i}_{2:K}),t}^{\mathsf{t}}X_{1,(j,\mathbf{i}_{2:K}),t}^{\mathsf{t}}-\mathsf{E}(X_{1,(i,\mathbf{i}_{2:K}),t}^{\mathsf{t}}X_{1,(j,\mathbf{i}_{2:K}),t}^{\mathsf{t}})$ for $i,j\in[p_1]$. By Lemma B.2, we have $\mathbf{Y}=\{Y_{ij,\mathbf{i}_{2:K},t},\ i_l\in[p_l],\ 2\le l\le K,\ t\in[n]\}$, a strongly mixing random field with the mixing coefficient as in Assumption 5 (i). Then by Lemma B.3 and Assumption 5 (i), we have for some $\nu>2$,

$$\begin{aligned}&\left|\mathsf{Cov}(Y_{ij,\mathbf{i}_{2:K},t},Y_{ij,\mathbf{i}'_{2:K},u})\right|\\ &\le 8\|Y_{ij,\mathbf{i}_{2:K},t}\|_\nu\|Y_{ij,\mathbf{i}'_{2:K},u}\|_\nu\exp\left(-\frac{c_0(\nu-2)(|t-u|+\sum_{l=2}^K|i_l-i_l'|)}{K\nu}\right),\end{aligned}$$

from the observation that $\max(\max_{2\le l\le K}|i_l-i_l'|,|t-u|)\ge K^{-1}(\sum_{l=2}^K|i_l-i_l'|+|t-u|)$. Further,

$$\|Y_{ij,\mathbf{i}_{2:K},t}\|_\nu^\nu \lesssim \tau^{2(\nu-1-\epsilon)}(|\mathcal{F}_t|_2^{2+2\epsilon}+\omega^{2+2\epsilon})$$

for all $i, j, \mathbf{i}_{2:K}$ and $t$, see (B.12) and Lemma B.1 (ii). Noticing that

$$T_{2,ij} = \frac{1}{np_{-1}} \sum_{t \in [n]} \otimes_{l=2}^K \sum_{i_l \in [p_l]} Y_{ij,\mathbf{i}_{2:K},t},$$

we have with $\nu = 2 + \log^{-1}(np)$,

$$\begin{aligned}
\|T_{2,ij}\|_2^2 &\le \frac{1}{(np_{-1})^2} \sum_{t,u \in [n]} \otimes_{l=2}^K \sum_{i_l, i_l' \in [p_l]} \mathsf{Cov}(Y_{ij,\mathbf{i}_{2:K},t}, Y_{ij,\mathbf{i}'_{2:K},u}) \\
&\lesssim \frac{1}{(np_{-1})^2} \sum_{t,u \in [n]} \otimes_{l=2}^K \sum_{i_l, i_l' \in [p_l]} \|Y_{ij,\mathbf{i}_{2:K},t}\|_\nu \|Y_{ij,\mathbf{i}'_{2:K},u}\|_\nu \\
&\qquad \times \exp\left( -\frac{c_0(\nu-2)(|t-u| + \sum_{l=2}^K |i_l - i_l'|)}{K\nu} \right) \\
&\lesssim \frac{\tau^{\frac{4(\nu-1-\epsilon)}{\nu}}}{np_{-1}} \cdot \frac{1}{n} \sum_{t,u \in [n]} (|\mathcal{F}_t|_2 + \omega)^{\frac{2+2\epsilon}{\nu}} (|\mathcal{F}_u|_2 + \omega)^{\frac{2+2\epsilon}{\nu}} \exp\left( -\frac{c_0(\nu-2)|t-u|}{K\nu} \right) \\
&\qquad \times \prod_{l=2}^K \frac{1}{p_l} \sum_{i_l, i_l' \in [p_l]} \exp\left( -\frac{c_0(\nu-2)|i_l - i_l'|}{K\nu} \right) \\
&\lesssim \frac{\tau^{2-2\epsilon}}{np_{-1}} \cdot \frac{1}{n} \sum_{t,u \in [n]} (|\mathcal{F}_t|_2 + \omega)^{1+\epsilon} (|\mathcal{F}_u|_2 + \omega)^{1+\epsilon} \exp\left( -\frac{c_0|t-u|}{3K\log(np)} \right) \\
&\qquad \times \prod_{l=2}^K \frac{1}{p_l} \sum_{i_l, i_l' \in [p_l]} \exp\left( -\frac{c_0|i_l - i_l'|}{3K\log(np)} \right) \lesssim \omega^{2+2\epsilon} \tau^{2-2\epsilon} c_\epsilon \frac{(K\log(np))^K}{np_{-1}},
\end{aligned}$$

where the final inequality follows from Assumption 5 (ii). The above upper bound does not depend on $i, j$ or $k$ and thus it follows that

$$\begin{aligned}
\frac{1}{p_k} \left\| \widehat{\mathbf{\Gamma}}^{(k)} - \mathbf{\Gamma}^{(k)} \right\|_F &\le \sqrt{\frac{1}{p_k^2} \sum_{i,j \in [p_k]} T_{1,ij}^2} + \frac{1}{p_k} \left\| \widehat{\mathbf{\Gamma}}^{(k)} - \mathsf{E}(\widehat{\mathbf{\Gamma}}^{(k)}) \right\|_F \\
&= O_P\left( \frac{\omega^{2+2\epsilon}}{\tau^{2\epsilon}} + \omega^{1+\epsilon}\tau^{1-\epsilon} \sqrt{\frac{\log^K(np_{-k})}{np_{-k}}} \right) = O_P\left( \omega^2 \left( \frac{\log^K(np)}{np_{-k}} \right)^{\frac{\epsilon}{1+\epsilon}} \right)
\end{aligned}$$

by Markov's inequality, with $\tau \asymp \tau_{n,p}^{(k)}$ in (10).

Finally, combining the bound on $p_k^{-1} \|\widehat{\mathbf{\Gamma}}^{(k)} - \mathbf{\Gamma}^{(k)}\|_F$ with Lemma A.2, the proof of the first claim is complete. □

*Proof of* (B.3). Let us set $\mathbf{S} = \mathbf{\Gamma}_\chi^{(k)}$, $\widetilde{\mathbf{S}} = \mathbf{\Gamma}^{(k)}$ and $\widehat{\mathbf{S}} = \widehat{\mathbf{\Gamma}}^{(k)}$. Then thanks to Lemma A.1, we have $\mu_{\chi,j}^{(k)}$, $j \in [r_k]$, fulfil the condition (C1) in Lemma B.4. Also, $\|\mathbf{\Gamma}^{(k)} - \mathbf{\Gamma}_\chi^{(k)}\| = \|\mathbf{\Gamma}_\xi^{(k)}\| \lesssim \omega^2$ from Lemma A.2. These establish that $\mathbf{\Gamma}_\chi^{(k)}$ and $\mathbf{\Gamma}^{(k)}$ meet (C3) in place of $\mathbf{S}$ and $\widetilde{\mathbf{S}}$.

Combining this with (B.2) (playing the role of (C2)), the claim follows from Lemma B.4 (i). □

### B.2.2 Supporting lemmas

**Lemma B.5.** *Let Assumptions 1, 2, 3, and 4 or 5 hold. For each $k \in [K]$, we have*

$$\left\| \left(p_k^{-1}\widehat{\mathbf{M}}_k\right)^{-1} - \left(p_k^{-1}\mathbf{M}_{\chi,k}\right)^{-1} \right\| = O_P\left(\psi_{n,p}^{(k)} \vee \frac{1}{p_k}\right),$$

*where $\mathbf{M}_{\chi,k} = diag(\mu_{\chi,1}^{(k)}, \ldots, \mu_{\chi,r_k}^{(k)})$.*

*Proof.* As noted in the proof of (B.3), $\mathbf{\Gamma}^{(k)}$ and $\mathbf{\Gamma}_\chi^{(k)}$ fulfil the conditions (C1) and (C3) in Lemma B.4 in place of $\widetilde{\mathbf{S}}$ and $\mathbf{S}$, respectively. The conclusions follow from (B.2) and Lemma B.4 (ii). □

**Lemma B.6.** *Let Assumptions 1, 2, 3 and 4 hold. For each $k \in [K]$, we have*

$$\frac{1}{np}\left\| \sum_{t\in[n]} \mathbf{Z}_{k,t}\mathbf{Z}_{k,t}^\top \right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p_k} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}}\right).$$

*Proof.* Note that

$$\frac{1}{n^2p^2}\left\| \sum_{t\in[n]} \mathbf{Z}_{k,t}\mathbf{Z}_{k,t}^\top \right\|^2 \le \frac{2}{n^2p^2}\left\| \sum_{t\in[n]} \left(\mathbf{Z}_{k,t}\mathbf{Z}_{k,t}^\top - \mathsf{E}\left(\mathbf{Z}_{k,t}\mathbf{Z}_{k,t}^\top\right)\right) \right\|^2 + \frac{2}{n^2p^2}\left\| \sum_{t\in[n]} \mathsf{E}\left(\mathbf{Z}_{k,t}\mathbf{Z}_{k,t}^\top\right) \right\|^2 =: U_1 + U_2.$$

Then with $\mathbf{Z}_{k,t} = [Z_{k,i\ell,t},\, i \in [p_k],\, \ell \in [p_{-k}]]$, we have

$$\begin{aligned}
\mathsf{E}(U_1) &\le \frac{2}{n^2p^2} \sum_{i,j\in[p_k]} \mathsf{E}\left[\left(\sum_{t\in[n]}\sum_{\ell\in[p_{-k}]} (Z_{k,i\ell,t}Z_{k,j\ell,t} - \mathsf{E}(Z_{k,i\ell,t}Z_{k,j\ell,t}))\right)^2\right] \\
&\le \frac{2}{n^2p^2} \sum_{i,j\in[p_k]} \sum_{\ell,m\in[p_{-k}]} \sum_{t,u\in[n]} |\mathsf{Cov}(Z_{k,i\ell,t}Z_{k,j\ell,t}, Z_{k,im,u}Z_{k,jm,u})| \\
&= \frac{2}{n^2p^2} \sum_{i,j\in[p_k]} \sum_{\ell\in[p_{-k}]} \sum_{t,u\in[n]} |\mathsf{Cov}(Z_{k,i\ell,t}Z_{k,j\ell,t}, Z_{k,i\ell,u}Z_{k,j\ell,u})| \\
&\le \frac{2}{n^2p^2} \sum_{i,j\in[p_k]} \sum_{\ell\in[p_{-k}]} \sum_{t,u\in[n]} \|Z_{k,i\ell,t}Z_{k,j\ell,t}\|_\nu \|Z_{k,i\ell,u}Z_{k,j\ell,u}\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right)
\end{aligned}$$

$$\lesssim \frac{(\tau_{n,p}^{(k)})^{\frac{4(\nu-1-\epsilon)}{\nu}}}{n^2p^2} \sum_{i\in[p_k]} \sum_{\ell\in[p_{-k}]} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^{\frac{2+2\epsilon}{\nu}} (|\mathcal{F}_u|_2+\omega)^{\frac{2+2\epsilon}{\nu}} \exp\left(-\frac{c_0|t-u|}{3\log(np_{-k})}\right)$$
$$+ \frac{1}{n^2p^2} \sum_{\substack{i,j\in[p_k]\\ i\neq j}} \sum_{\ell\in[p_{-k}]} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^2 (|\mathcal{F}_u|_2+\omega)^2 \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right)$$
$$\lesssim \frac{(\tau_{n,p}^{(k)})^{2-2\epsilon}}{np}\log(np_{-k}) + \frac{M_n^{2-2\epsilon}}{np_{-k}} \lesssim \left(\frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} + \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right)^2$$

with $\nu \in \{2+\log^{-1}(np_{-k}), 2+2\epsilon\}$ for the case of $i=j$ and $i\neq j$, respectively, the first equality is due to the cross-sectional independence (Assumption 4 (i)), the third inequality holds due to Lemmas B.2 and B.3 and Assumption 4 (i), the fourth due to Lemma B.1 (i) and (ii), and the penultimate one follows from Assumptions 3 (iii) and 4 (ii). As for $U_2$, notice that $\sum_{t\in[n]} \mathsf{E}(\mathbf{Z}_{k,t}\mathbf{Z}_{k,t}^\top)$ is a diagonal matrix such that

$$U_2 \lesssim \max_{i\in[p_k]} \left(\frac{1}{np}\sum_{t\in[n]}\sum_{\ell\in[p_{-k}]} \mathsf{E}(Z_{k,i\ell,t}^2)\right)^2 \lesssim \max_{i\in[p_k]} \left(\frac{1}{np}\sum_{t\in[n]}\sum_{\ell\in[p_{-k}]} (|\mathcal{F}_t|_2^2+\omega^2)\right)^2 \lesssim \left(\frac{\omega^2}{p_k}\right)^2$$

by Lemma B.1 (i) and Assumption 3 (ii). Collecting the bounds on $U_1$ and $U_2$ and by Markov's inequality, the conclusion follows. □

**Lemma B.7.** *Let Assumptions 1, 2, 3 and 4 hold. For each $k\in[K]$, we have*

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right),$$
$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})\mathbf{Z}_{k,t}^\top\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right).$$

*Proof.* With $\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}}) = [\mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}),\ i\in[p_k],\ \ell\in[p_{-k}]]$, we have

$$\mathsf{E}\left(\left\|\frac{1}{np}\sum_{t\in[n]} \mathbf{Z}_{k,t}\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\|^2\right)$$
$$\leq \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\sum_{t,u\in[n]} \mathsf{E}(X_{k,j\ell,t}^{\mathsf{t}})\mathsf{E}(X_{k,jm,u}^{\mathsf{t}})\mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,im,u})$$
$$= \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell\in[p_{-k}]}\sum_{t,u\in[n]} \mathsf{E}(X_{k,j\ell,t}^{\mathsf{t}})\mathsf{E}(X_{k,j\ell,u}^{\mathsf{t}})\mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,i\ell,u})$$
$$\leq \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell\in[p_{-k}]}\sum_{t,u\in[n]} \left|\mathsf{E}(X_{k,j\ell,t}^{\mathsf{t}})\mathsf{E}(X_{k,j\ell,u}^{\mathsf{t}})\right| \|Z_{k,i\ell,t}\|_\nu \|Z_{k,i\ell,u}\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right)$$

$$\lesssim \frac{1}{n^2p^2} \sum_{i,j\in[p_k]} \sum_{\ell\in[p_{-k}]} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^2(|\mathcal{F}_u|_2+\omega)^2 \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \lesssim \frac{M_n^{2-2\epsilon}}{np_{-k}}$$

with $\nu = 2+2\epsilon$, where the first equality is due to the cross-sectional independence (Assumption 4 (i)), the second inequality holds due to Lemmas B.2 and B.3 and Assumption 4 (i), the third due to Lemma B.1 (i), and the last one follows from Assumptions 3 (iii) and 4 (ii). The second result follows by the analogous arguments. □

**Lemma B.8.** *Let Assumptions 1, 2, 3, and 4 or 5 hold. For each $k \in [K]$, we have*

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}}-\mathbf{X}_t)^\top\right\| = O_P\left(\frac{M_n^\epsilon\sqrt{\log(np_{-k})}}{(\tau_{n,p}^{(k)})^\epsilon}\cdot\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right),$$
$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}}-\mathbf{X}_t)\mathsf{E}(\mathbf{X}_{k,t})^\top\right\| = O_P\left(\frac{M_n^\epsilon\sqrt{\log(np_{-k})}}{(\tau_{n,p}^{(k)})^\epsilon}\cdot\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right).$$

*Proof.* With $\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}}-\mathbf{X}_{k,t}) = [\mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}-X_{k,i\ell,t}),\ i\in[p_k],\ \ell\in[p_{-k}]]$, we have

$$\begin{aligned}
&\left\|\frac{1}{np}\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}}-\mathbf{X}_t)^\top\right\|_F^2 \\
&= \sum_{i,j\in[p_k]} \frac{1}{n^2p^2} \sum_{\ell,m\in[p_{-k}]} \sum_{t,u\in[n]} \mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}})\mathsf{E}(X_{k,im,u}^{\mathsf{t}})\mathsf{E}(X_{k,j\ell,t}^{\mathsf{t}}-X_{k,j\ell,t})\mathsf{E}(X_{k,jm,u}^{\mathsf{t}}-X_{k,jm,u}) \\
&\lesssim \sum_{i,j\in[p_k]} \frac{1}{n^2p^2} \sum_{\ell,m\in[p_{-k}]} \sum_{t,u\in[n]} \frac{(|\mathcal{F}_t|_2+\omega)^{3+2\epsilon}(|\mathcal{F}_u|_2+\omega)^{3+2\epsilon}}{(\tau_{n,p}^{(k)})^{2+4\epsilon}} \\
&\lesssim \frac{\omega^{4+4\epsilon}M_n^2}{(\tau_{n,p}^{(k)})^{2+4\epsilon}} \lesssim \left(\frac{M_n^\epsilon\sqrt{\log(np_{-k})}}{(\tau_{n,p}^{(k)})^\epsilon}\cdot\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right)^2,
\end{aligned}$$

where the first inequality is due to Lemma B.1 (i) and (iii) and the second from Assumption 3 (ii) and (iii). The second result follows by the analogous arguments. □

**Lemma B.9.** *Let Assumptions 1 and 2 hold. For each $k \in [K]$, there exists a matrix $\mathbf{H}_k \in \mathbb{R}^{r_k\times r_k}$ satisfying $\mathbf{H}_k^\top\mathbf{H}_k = \mathbf{I}_{r_k}$, such that*

$$\frac{1}{\sqrt{p_k}}\mathbf{\Lambda}_k\mathbf{H}_k = \mathbf{E}_{\chi,k}.$$

*Proof.* By definition,

$$\mathbf{E}_{\chi,k}\mathbf{M}_{\chi,k} = \mathbf{\Gamma}_\chi^{(k)}\mathbf{E}_{\chi,k} = \mathbf{\Lambda}_k \underbrace{\left(\frac{1}{n}\sum_{t\in[n]} \mathsf{mat}_k(\mathcal{F}_t)\mathsf{mat}_k(\mathcal{F}_t)^\top\right)}_{=:\bar{\mathbf{\Gamma}}_f^{(k)}} \mathbf{\Lambda}_k^\top \mathbf{E}_{\chi,k},$$

where $\|\bar{\mathbf{\Gamma}}_f^{(k)}\| = O(1)$ under Assumptions 1 and 2. Let us set

$$\mathbf{H}_k = \sqrt{p_k}\bar{\mathbf{\Gamma}}_f^{(k)}\mathbf{\Lambda}_k^\top\mathbf{E}_{\chi,k}(\mathbf{M}_{\chi,k})^{-1}.$$

Then, we have

$$\begin{aligned}
\mathbf{H}_k^\top\mathbf{H}_k &= p_k(\mathbf{M}_{\chi,k})^{-1}\mathbf{E}_{\chi,k}^\top\mathbf{\Lambda}_k\left(\bar{\mathbf{\Gamma}}_f^{(k)}\right)^2\mathbf{\Lambda}_k^\top\mathbf{E}_{\chi,k}(\mathbf{M}_{\chi,k})^{-1} \\
&= (\mathbf{M}_{\chi,k})^{-1}\mathbf{E}_{\chi,k}^\top\mathbf{\Lambda}_k\bar{\mathbf{\Gamma}}_f^{(k)}\mathbf{\Lambda}_k^\top\mathbf{\Lambda}_k\bar{\mathbf{\Gamma}}_f^{(k)}\mathbf{\Lambda}_k^\top\mathbf{E}_{\chi,k}(\mathbf{M}_{\chi,k})^{-1} = \mathbf{I}_{r_k}.
\end{aligned}$$

□

### B.3 Proof of Theorem 2

Throughout, we suppress the dependence on $\tau$ where there is no confusion. For each $k \in [K]$, we decompose $\check{\mathbf{\Gamma}}^{(k),[1]}$ as

$$\begin{aligned}
\frac{1}{p_k}\check{\mathbf{\Gamma}}^{(k),[1]} &= \frac{1}{np}\sum_{t\in[n]} \mathbf{X}_{k,t}^{\mathsf{t}}\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top(\mathbf{X}_{k,t}^{\mathsf{t}})^\top \\
&= \frac{1}{np}\sum_{t\in[n]} \mathbf{Z}_{k,t}\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathbf{Z}_{k,t}^\top \\
&\quad + \frac{1}{np}\sum_{t\in[n]} \mathbf{Z}_{k,t}\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top + \frac{1}{np}\sum_{t\in[n]}\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathbf{Z}_{k,t}^\top \\
&\quad + \frac{1}{np}\sum_{t\in[n]}\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t}\right)\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top \\
&\quad + \frac{1}{np}\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,t})\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t}\right)^\top \\
&\quad + \frac{1}{np}\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,t})\left(\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}}\mathbf{\Delta}_k\mathbf{\Delta}_k^\top\right)\mathsf{E}(\mathbf{X}_{k,t})^\top \\
&\quad + \frac{1}{np_kp_{-k}^2}\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,t})\mathbf{\Delta}_k\mathbf{\Delta}_k^\top\mathsf{E}(\mathbf{X}_{k,t})^\top \\
&=: T_1 + T_{2,1} + T_{2,2} + T_{3,1} + T_{3,2} + T_4 + \frac{1}{p_k}\mathbf{\Gamma}_\chi^{(k)}.
\end{aligned} \tag{B.13}$$

where $\mathbf{\Gamma}_\chi^{(k)}$ is defined in (A.1). Based on this, we derive the following rate of estimation for $\breve{\mathbf{\Gamma}}^{(k),[1]}(\tau)$:

**Proposition B.2.** *Let Assumptions 1, 2 and 3 hold. Then for each $k \in [K]$,*

$$\frac{1}{p_k}\left\|\breve{\mathbf{\Gamma}}^{(k),[1]}(\tau) - \mathbf{\Gamma}_\chi^{(k)}\right\| = \begin{cases} O_P\left[\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \bar{\psi}_{n,p}^{(k)}\left(\frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} + \frac{M_n^{1-\epsilon}}{\sqrt{n}}\right) \vee \frac{\bar{\psi}_{n,p}}{\sqrt{p}} \vee \sum_{k' \in [K]\setminus\{k\}}\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p_{k'}} \vee \frac{\psi_{n,p}^{(k')}}{\sqrt{p_{k'}}}\right)\right] \\ \qquad \text{under Assumption 4,} \\ O_P\left(\bar{\psi}_{n,p} \vee \sum_{k' \in [K]\setminus\{k\}} \frac{1}{p_{k'}}\right) \\ \qquad \text{under Assumption 5,} \end{cases}$$

*with $\tau = \tau^{(k)}$ chosen as in* (10).

*Proof.* Noting the decomposition in (B.13), we obtain the desired rates by collecting the bounds on $T_1$–$T_4$ derived in Lemmas B.12–B.15. □

For each $k \in [K]$, denote by $\breve{\mathbf{M}}_k^{[1]}(\tau) \in \mathbb{R}^{r_k \times r_k}$ the diagonal matrix containing the eigenvalues $\breve{\mu}_j^{(k),[1]}(\tau)$, $j \in [r_k]$, of $\breve{\mathbf{\Gamma}}^{(k),[1]}(\tau)$ on its diagonal. By Proposition B.2, Lemma B.16 and the arguments analogous to those adopted in proving (B.4) and (B.5), we have $p_k^{-1}\breve{\mathbf{M}}_k^{[1]}$ asymptotically invertible and

$$\left\|\left(p_k^{-1}\breve{\mathbf{M}}_k^{[1]}\right)^{-1}\right\| \le \frac{1}{\alpha_{r_k}^{(k)}} + o_P(1) = O_P(1). \tag{B.14}$$

Let us set

$$\breve{\mathbf{H}}_k^{[1]} = \frac{1}{n\sqrt{p_k}p_{-k}} \sum_{t \in [n]} \mathsf{mat}_k(\mathcal{F}_t)\mathbf{\Delta}_k^\top \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top \mathbf{\Delta}_k \mathsf{mat}_k(\mathcal{F}_t)^\top \mathbf{\Lambda}_k^\top \breve{\mathbf{E}}_k^{[1]} \left(\frac{1}{p_k}\breve{\mathbf{M}}_k^{[1]}\right)^{-1},$$

such that

$$\frac{1}{\sqrt{p_k}}\mathbf{\Lambda}_k \breve{\mathbf{H}}_k^{[1]} = \frac{1}{np}\sum_{t \in [n]} \mathsf{E}(\mathbf{X}_{k,t})\widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top \mathsf{E}(\mathbf{X}_{k,t})^\top \breve{\mathbf{E}}_k^{[1]} \left(\frac{1}{p_k}\breve{\mathbf{M}}_k^{[1]}\right)^{-1}.$$

Then noting that $\breve{\mathbf{\Gamma}}^{(k),[1]}\breve{\mathbf{E}}_k^{[1]} = \breve{\mathbf{E}}_k^{[1]}\breve{\mathbf{M}}_k^{[1]}$, we have

$$\breve{\mathbf{E}}_k^{[1]} - \frac{1}{\sqrt{p_k}}\mathbf{\Lambda}_k \breve{\mathbf{H}}_k^{[1]} = (T_1 + T_{2,1} + T_{2,2} + T_{3,1} + T_{3,2})\,\breve{\mathbf{E}}_k^{[1]} \left(\frac{1}{p_k}\breve{\mathbf{M}}_k^{[1]}\right)^{-1}.$$

By Lemmas B.12–B.14 and (B.14),

$$\left\|\breve{\mathbf{E}}_k^{[1]} - \frac{1}{\sqrt{p_k}}\mathbf{\Lambda}_k\breve{\mathbf{H}}_k^{[1]}\right\| = \begin{cases} O_P\left[\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p} \vee \bar{\psi}_{n,p}^{(k)}\left(\frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} + \frac{M_n^{1-\epsilon}}{\sqrt{n}}\right) \vee \frac{\bar{\psi}_{n,p}}{\sqrt{p}}\right] \\ \quad \text{under Assumption 4,} \\ O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p} \vee \psi_{n,p}^{(k)} \vee \frac{M_n^{1-\epsilon}\bar{\psi}_{n,p}^{(k)}}{\sqrt{n}}\right) \\ \quad \text{under Assumption 5.} \end{cases}$$

We conclude the proof by noting that

$$\|\breve{\mathbf{H}}_k^{[1]}\| \le \frac{1}{n\sqrt{p_k}p_{-k}} \sum_{t\in[n]} |\mathcal{F}_t|_2^2 \|\mathbf{\Delta}_k\|^2 \|\mathbf{\Lambda}_k\| \left\|\left(\frac{1}{p_k}\breve{\mathbf{M}}_k^{[1]}\right)^{-1}\right\| \le (\alpha_{r_k}^{(k)})^{-1}\omega^2(1+o_P(1))$$

by Assumptions 1 and 3 (ii) and (B.14), and hence

$$\mathbf{I}_{r_k} = (\breve{\mathbf{E}}_k^{[1]})^\top \breve{\mathbf{E}}_k^{[1]} = \frac{1}{p_k}(\breve{\mathbf{H}}_k^{[1]})^\top \mathbf{\Lambda}_k^\top \mathbf{\Lambda}_k \breve{\mathbf{H}}_k^{[1]} + o_P(1) = (\breve{\mathbf{H}}_k^{[1]})^\top \breve{\mathbf{H}}_k^{[1]} + o_P(1).$$

### B.3.1 Supporting lemmas

Throughout, we suppress the dependence on $\tau$ where there is no confusion.

**Lemma B.10.** *Let Assumptions 1, 2, 3, and 4 or 5 hold, and define*

$$\mathbf{D}_k = \mathbf{E}_K \otimes \cdots \otimes \mathbf{E}_{k+1} \otimes \mathbf{E}_{k-1} \otimes \cdots \otimes \mathbf{E}_1.$$

*Recall $\mathbf{J}_k$ and $\widehat{\mathbf{J}}_k$ from Proposition B.1 and $\widehat{\mathbf{H}}_k$ from Theorem 1. Then for all $k \in [K]$, with $\tau = \tau^{(k)}$ chosen as in* (10)*, we have:*

*(i) Letting $\mathbf{J}_{-k} = \mathbf{J}_K \otimes \cdots \otimes \mathbf{J}_{k+1} \otimes \mathbf{J}_{k-1} \otimes \cdots \otimes \mathbf{J}_1$,*

$$\left\|\widehat{\mathbf{D}}_k - \mathbf{D}_k\mathbf{J}_{-k}\right\| = O_P\left(\bar{\psi}_{n,p}^{(k)}\right).$$

*(ii) Letting $\widehat{\mathbf{J}}_{-k} = \widehat{\mathbf{J}}_{k+1} \otimes \widehat{\mathbf{J}}_{k-1} \otimes \cdots \otimes \widehat{\mathbf{J}}_1$,*

$$\left\|\widehat{\mathbf{D}}_k - \frac{1}{\sqrt{p_{-k}}}\mathbf{\Delta}_k\widehat{\mathbf{J}}_{-k}\right\| = O_P\left(\sum_{k'\in[K]\setminus\{k\}}\left(\psi_{n,p}^{(k')} \vee \frac{1}{p_{k'}}\right)\right).$$

*(iii) Suppose that Assumption 4 holds. Letting* $\widehat{\mathbf{H}}_{-k} = \widehat{\mathbf{H}}_K \otimes \cdots \otimes \widehat{\mathbf{H}}_{k+1} \otimes \widehat{\mathbf{H}}_{k-1} \otimes \cdots \otimes \widehat{\mathbf{H}}_1$,

$$\left\| \widehat{\mathbf{D}}_k - \frac{1}{\sqrt{p_{-k}}} \boldsymbol{\Delta}_k \widehat{\mathbf{H}}_{-k} \right\| = O_P \left( \sum_{k' \in [K] \setminus \{k\}} \left( \frac{\psi_{n,p}^{(k')}}{\sqrt{p_{k'}}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p_{k'}} \right) \right).$$

*(iv) There exists some constant* $C > 0$ *such that* $|\mathbf{D}_k|_\infty \le C p_{-k}^{-1/2}$ *and* $|\boldsymbol{\Delta}_k|_\infty \le C$.

*Proof.* WLOG, we consider the case where $k = 1$. Note that,

$$\begin{aligned} \widehat{\mathbf{D}}_1 - \mathbf{D}_1 \mathbf{J}_{-1} = & (\widehat{\mathbf{E}}_K - \mathbf{E}_K \mathbf{J}_K) \otimes_{k'=K-1}^{2} \widehat{\mathbf{E}}_{k'} + (\mathbf{E}_K \mathbf{J}_K) \otimes (\widehat{\mathbf{E}}_{K-1} - \mathbf{E}_{K-1} \mathbf{J}_{K-1}) \otimes_{k'=K-2}^{2} \widehat{\mathbf{E}}_{k'} \\ & + \ldots + \otimes_{k'=K}^{3} \mathbf{E}_k \mathbf{J}_k \otimes (\widehat{\mathbf{E}}_2 - \mathbf{E}_2 \mathbf{J}_2) \end{aligned}$$

such that

$$\begin{aligned} \left\| \widehat{\mathbf{D}}_1 - \mathbf{D}_1 \mathbf{J}_{-1} \right\| \le & \left\| \widehat{\mathbf{E}}_K - \mathbf{E}_K \mathbf{J}_K \right\| \prod_{k'=K-1}^{2} \left\| \widehat{\mathbf{E}}_{k'} \right\| \\ & + \| \mathbf{E}_K \mathbf{J}_K \| \left\| \widehat{\mathbf{E}}_{K-1} - \mathbf{E}_{K-1} \mathbf{J}_{K-1} \right\| \prod_{k'=K-2}^{2} \left\| \widehat{\mathbf{E}}_{k'} \right\| \\ & + \ldots + \prod_{k'=K}^{3} \| \mathbf{E}_k \mathbf{J}_k \| \left\| \widehat{\mathbf{E}}_2 - \mathbf{E}_2 \mathbf{J}_2 \right\| = O_P \left( \sum_{k'=2}^{K} \psi_{n,p}^{(k')} \right) \end{aligned}$$

by (B.3) in Proposition B.1, which proves (i). The proofs of (ii) and (iii) take the analogous steps, the former utilising (B.3) and the latter (B.8), and thus is omitted. For (iv), observe that by Lemma B.1 (i), Assumption 3 and Cauchy-Schwarz inequality, $|\boldsymbol{\Gamma}^{(k)}|_\infty \lesssim n^{-1} \sum_{t \in [n]} (|\mathcal{F}_t|_2^2 + \omega^2) \le 2\omega^2$. Combining this with the arguments used in the proof of (B.3), by Lemma B.4 (iii), we have

$$|\mathbf{E}_k|_\infty \le \frac{C'}{\sqrt{p_k}}$$

for all $k \in [K]$ and some constant $C' > 0$, which proves the first claim with $C = (C')^{K-1}$. The second claim follows with $C = \bar{\lambda}^{K-1}$ due to Assumption 1 (ii). □

**Lemma B.11.** *Let Assumptions 1, 2, 3, and 4 or 5 hold. Then for all* $k \in [K]$,

$$\left\| \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \mathbf{D}_k \mathbf{D}_k^\top \right\| = O_P \left( \bar{\psi}_{n,p}^{(k)} \right). \tag{B.15}$$

*Also, under Assumption 4,*

$$\left\| \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right\| = O_P \left( \sum_{k' \in [K] \setminus \{k\}} \left( \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p_{k'}} \vee \frac{\psi_{n,p}^{(k')}}{\sqrt{p_{k'}}} \right) \right), \tag{B.16}$$

*while under Assumption 5,*

$$\left\|\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right\| = O_P\left(\sum_{k'\in[K]\setminus\{k\}}\left(\psi_{n,p}^{(k')}\vee\frac{1}{p_{k'}}\right)\right). \tag{B.17}$$

*Proof.* By Lemma B.10 (i), we have

$$\left\|\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top - \mathbf{D}_k\mathbf{D}_k^\top\right\| \le \left\|\widehat{\mathbf{D}}_k\left(\widehat{\mathbf{D}}_k - \mathbf{D}_k\mathbf{J}_{-k}\right)^\top\right\| + \left\|\left(\widehat{\mathbf{D}}_k - \mathbf{D}_k\mathbf{J}_{-k}\right)\mathbf{J}_{-k}\mathbf{D}_k^\top\right\| = O_P\left(\bar{\psi}_{n,p}^{(k)}\right),$$

which proves (B.15). Similarly under Assumption 5,

$$\begin{aligned}\left\|\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right\| &\le \left\|\widehat{\mathbf{D}}_k\left(\widehat{\mathbf{D}}_k - \frac{1}{\sqrt{p_{-k}}}\boldsymbol{\Delta}_k\widehat{\mathbf{J}}_{-k}\right)^\top\right\| \\ &\quad + \left\|\widehat{\mathbf{D}}_k - \frac{1}{\sqrt{p_{-k}}}\boldsymbol{\Delta}_k\widehat{\mathbf{J}}_{-k}\right\|\left\|\widehat{\mathbf{J}}_{-k}\right\|\left\|\frac{1}{\sqrt{p_{-k}}}\boldsymbol{\Delta}_k\right\| \\ &= O_P\left(\sum_{k'\in[K]\setminus\{k\}}\left(\psi_{n,p}^{(k')}\vee\frac{1}{p_{k'}}\right)\right)\end{aligned}$$

which follows from Lemma B.10 (ii), Proposition B.1 and Assumption 1 (i), thus proving (B.17).

Recall that $\widehat{\mathbf{H}}_k \in \mathbb{R}^{r_k\times r_k}$ is asymptotically invertible with $\|\widehat{\mathbf{H}}_k\| = O_P(1)$. Further, under Assumption 4, by Theorem 1, we have

$$\mathbf{I}_{r_k} = \widehat{\mathbf{E}}_k^\top\widehat{\mathbf{E}}_k = \widehat{\mathbf{H}}_k^\top\widehat{\mathbf{H}}_k + O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\vee\frac{1}{p_k}\vee\frac{\psi_{n,p}^{(k)}}{\sqrt{p}_k}\right),$$

taking steps analogous to those used in (B.9). From the above, it follows that $\|\widehat{\mathbf{H}}_k^{-1}\| = O_P(1)$ and by the same token, we have $\|\widehat{\mathbf{H}}_{-k}^{-1}\| = O_P(1)$. Then, we have

$$\begin{aligned}&\left\|\widehat{\mathbf{D}}_k\widehat{\mathbf{H}}_{-k}^\top - \frac{1}{\sqrt{p_{-k}}}\boldsymbol{\Delta}_k\right\| = \left\|\left(\widehat{\mathbf{D}}_k\widehat{\mathbf{H}}_{-k}^\top\widehat{\mathbf{H}}_{-k} - \frac{1}{\sqrt{p_{-k}}}\boldsymbol{\Delta}_k\widehat{\mathbf{H}}_{-k}\right)\widehat{\mathbf{H}}_{-k}^{-1}\right\| \\ &\le \left\|\widehat{\mathbf{D}}_k\left\{\mathbf{I}_{r_{-k}} + O_P\left[\sum_{k'\in[K]\setminus\{k\}}\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}}\vee\frac{1}{p_{k'}}\vee\frac{\psi_{n,p}^{(k')}}{\sqrt{p}_{k'}}\right)\right]\right\} - \frac{1}{\sqrt{p_{-k}}}\boldsymbol{\Delta}_k\widehat{\mathbf{H}}_{-k}\right\|\left\|\widehat{\mathbf{H}}_{-k}^{-1}\right\| \\ &= O_P\left(\sum_{k'\in[K]\setminus\{k\}}\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}}\vee\frac{1}{p_{k'}}\vee\frac{\psi_{n,p}^{(k')}}{\sqrt{p}_{k'}}\right)\right)\end{aligned}$$

by Lemmas B.10 (iii). Then,

$$\left\|\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right\|$$

$$\le \left\| \widehat{\mathbf{D}}_k \left( \widehat{\mathbf{D}}_k - \frac{1}{\sqrt{p_{-k}}} \mathbf{\Delta}_k \widehat{\mathbf{H}}_k \right)^\top \right\| + \left\| \left( \widehat{\mathbf{D}}_k \widehat{\mathbf{H}}_k^\top - \frac{1}{\sqrt{p_{-k}}} \mathbf{\Delta}_k \right) \frac{1}{\sqrt{p_{-k}}} \mathbf{\Delta}_k^\top \right\|$$
$$= O_P \left( \sum_{k' \in [K] \setminus \{k\}} \left( \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p_{k'}} \vee \frac{\psi_{n,p}^{(k')}}{\sqrt{p_{k'}}} \right) \right),$$

which proves (B.16). □

Lemmas B.12–B.15 analyse the terms involved in (B.13).

**Lemma B.12.** *Let Assumptions 1, 2 and 3 hold. For each $k \in [K]$, we have the followings:*

*(i) Under Assumption 4,*

$$\frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top \mathbf{Z}_{k,t}^\top \right\| = O_P \left[ \frac{\psi_{n,p}^{(k)} \bar{\psi}_{n,p}^{(k)}}{\sqrt{p_k}} \vee \frac{\bar{\psi}_{n,p}}{\sqrt{p}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{n}} \left( \frac{1}{p_{-k}} \vee \bar{\psi}_{n,p}^{(k)} \right) \vee \frac{1}{p} \right].$$

*(ii) Under Assumption 5,*

$$\frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top \mathbf{Z}_{k,t}^\top \right\| = O_P \left( \psi_{n,p}^{(k)} \vee \frac{1}{p} \right).$$

*Proof of Lemma B.12 (i).* Let us write

$$\frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top \mathbf{Z}_{k,t}^\top \right\| \le \frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \left( \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \mathbf{D}_k \mathbf{D}_k^\top \right) \mathbf{Z}_{k,t}^\top \right\|$$
$$+ \frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \mathbf{D}_k \mathbf{D}_k^\top \mathbf{Z}_{k,t}^\top \right\| =: U_1 + U_2. \tag{B.18}$$

By Lemma B.17,

$$U_1^2 \le \underbrace{\frac{1}{n^2 p^2} \sum_{i,j \in [p_k]} \sum_{\ell, m \in [p_{-k}]} \left( \sum_{t \in [n]} Z_{k,i\ell,t} Z_{k,jm,t} \right)^2}_{V_1} \left\| \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \mathbf{D}_k \mathbf{D}_k^\top \right\|^2, \tag{B.19}$$

where

$$\mathsf{E}(V_1) \le \frac{2}{n^2 p^2} \sum_{i,j \in [p_k]} \sum_{\ell, m \in [p_{-k}]} \sum_{t,u \in [n]} \mathsf{Cov}(Z_{k,i\ell,t} Z_{k,jm,t}, Z_{k,i\ell,u} Z_{k,jm,u})$$
$$+ \frac{2}{n^2 p^2} \sum_{i,j \in [p_k]} \sum_{\ell, m \in [p_{-k}]} \left( \sum_{t \in [n]} \mathsf{E}(Z_{k,i\ell,t} Z_{k,jm,t}) \right)^2 =: V_{1,1} + V_{1,2}.$$

From Lemmas B.1 (i) and (ii), B.2 and B.3 and Assumptions 3 (iii) and 4,

$$\begin{aligned}
V_{1,1} &\lesssim \frac{2}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\sum_{t,u\in[n]}\|Z_{k,i\ell,t}Z_{k,jm,t}\|_\nu\|Z_{k,i\ell,u}Z_{k,jm,u}\|_\nu\exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right)\\
&\le \frac{2}{n^2p^2}\sum_{i\in[p_k]}\sum_{\ell\in[p_{-k}]}\sum_{t,u\in[n]}\|Z^2_{k,i\ell,t}\|_\nu\|Z^2_{k,i\ell,u}\|_\nu\exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right)\\
&\quad+\frac{2}{n^2p^2}\sum_{\substack{i,j\in[p_k]\\ i\neq j}}\sum_{\substack{\ell,m\in[p_{-k}]\\ \ell\neq m}}\sum_{t,u\in[n]}\|Z_{k,i\ell,t}\|_\nu\|Z_{k,jm,t}\|_\nu\\
&\qquad\qquad\qquad\times\|Z_{k,i\ell,u}\|_\nu\|Z_{k,jm,u}\|_\nu\exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right)\\
&\lesssim \frac{(\tau^{(k)}_{n,p})^{\frac{4(\nu-1-\epsilon)}{\nu}}}{n^2p}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^{\frac{2+2\epsilon}{\nu}}(|\mathcal{F}_u|_2+\omega)^{\frac{2+2\epsilon}{\nu}}\exp\left(-\frac{c_0|t-u|}{3\log(np_{-k})}\right)\\
&\quad+\frac{1}{n^2}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^2(|\mathcal{F}_u|_2+\omega)^2\exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right)\\
&\lesssim \frac{c_\epsilon(\tau^{(k)}_{n,p})^{2-2\epsilon}}{p_k}\cdot\frac{\log(np_{-k})}{np_{-k}}+\frac{M_n^{2-2\epsilon}c_\epsilon}{n}\lesssim\left(\frac{\psi^{(k)}_{n,p}}{\sqrt{p_k}}+\frac{M_n^{1-\epsilon}}{\sqrt{n}}\right)^2
\end{aligned}$$

with $\nu\in\{2+\log^{-1}(np_{-k}),2+2\epsilon\}$ for the case of $i=j$ and $i\neq j$, respectively. Also by Lemma B.1 (i) and Assumptions 3 (ii) and 4 (i), we have

$$V_{1,2}=\frac{2}{n^2p^2}\sum_{i\in[p_k]}\sum_{\ell\in[p_{-k}]}\left(\sum_{t\in[n]}\mathsf{E}(Z^2_{k,i\ell,t})\right)^2\lesssim\frac{1}{p}\left(\frac{1}{n}\sum_{t\in[n]}(|\mathcal{F}_t|_2^2+\omega^2)\right)^2\lesssim\frac{\omega^4}{p}.$$

Combining the bounds on $V_{1,1}$ and $V_{1,2}$ with (B.15) and (B.19), by Markov's inequality,

$$U_1=O_P\left[\left(\frac{\psi^{(k)}_{n,p}}{\sqrt{p_k}}\vee\frac{M_n^{1-\epsilon}}{\sqrt{n}}\vee\frac{\omega^2}{\sqrt{p}}\right)\bar{\psi}^{(k)}_{n,p}\right].$$

As for $U_2$, writing $\mathbf{D}_k=[d^{(k)}_{\ell q},\ \ell\in[p_{-k}],\ q\in[r_{-k}]]$,

$$\begin{aligned}
\mathsf{E}(U_2^2)\le{}&\frac{2}{n^2p^2}\sum_{i,j\in[p_k]}\mathsf{E}\left[\left(\sum_{t\in[n]}\sum_{\ell,m\in[p_{-k}]}\sum_{q\in[r_{-k}]}d^{(k)}_{\ell q}d^{(k)}_{mq}\left(Z_{k,i\ell,t}Z_{k,jm,t}-\mathsf{E}(Z_{k,i\ell,t}Z_{k,jm,t})\right)\right)^2\right]\\
&+\frac{2}{n^2p^2}\left\|\sum_{t\in[n]}\mathsf{E}\left(\mathbf{Z}_{k,t}\mathbf{D}_k\mathbf{D}_k^\top\mathbf{Z}_{k,t}^\top\right)\right\|^2=:V_{2,1}+V_{2,2}.
\end{aligned}$$

Then, by Assumptions 3 (iii) and 4, Lemmas B.1 (i) and (ii), B.2, B.3 and B.10 (iv),

$$
\begin{aligned}
V_{2,1} &\lesssim \frac{1}{n^2p^2} \sum_{i,j\in[p_k]} \sum_{\ell,\ell',m,m'\in[p_{-k}]} \sum_{q,q'\in[r_{-k}]} \sum_{t,u\in[n]} d_{\ell q}^{(k)} d_{mq}^{(k)} d_{\ell' q'}^{(k)} d_{m'q'}^{(k)} \mathsf{Cov}(Z_{k,i\ell,t}Z_{k,jm,t}, Z_{k,i\ell',u}Z_{k,jm',u}) \\
&\le \frac{1}{n^2p^2} \sum_{i\in[p_k]} \sum_{\substack{\ell,\ell',m,m'\in[p_{-k}] \\ (\ell,m)=(\ell',m') \text{ or} \\ (\ell,m)=(m',\ell') \text{ or} \\ (\ell,\ell')=(m,m')}} \sum_{q,q'\in[r_{-k}]} \sum_{t,u\in[n]} d_{\ell q}^{(k)} d_{mq}^{(k)} d_{\ell' q'}^{(k)} d_{m'q'}^{(k)} \|Z_{k,i\ell,t}Z_{k,im,t}\|_\nu \|Z_{k,i\ell',u}Z_{k,im',u}\|_\nu \\
&\quad \times \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right) + \frac{1}{n^2p^2} \sum_{\substack{i,j\in[p_k] \\ i\neq j}} \sum_{\ell,m\in[p_{-k}]} \sum_{q,q'\in[r_{-k}]} \sum_{t,u\in[n]} d_{\ell q}^{(k)} d_{mq}^{(k)} d_{\ell q'}^{(k)} d_{mq'}^{(k)} \\
&\quad \times \|Z_{k,i\ell,t}\|_\nu \|Z_{k,jm,t}\|_\nu \|Z_{k,i\ell,u}\|_\nu \|Z_{k,jm,u}\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right) \\
&\lesssim \frac{(\tau_{n,p}^{(k)})^{2-2\epsilon}}{n^2p_kp_{-k}^2} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^{1+\epsilon}(|\mathcal{F}_u|_2+\omega)^{1+\epsilon} \exp\left(-\frac{c_0|t-u|}{3\log(np_{-k})}\right) \\
&\quad + \frac{M_n^{2-2\epsilon}}{n^2p_{-k}^2} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^{1+\epsilon}(|\mathcal{F}_u|_2+\omega)^{1+\epsilon} \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \\
&\lesssim \left(\frac{\psi_{n,p}^{(k)}}{\sqrt{p}} + \frac{M_n^{1-\epsilon}}{\sqrt{n}p_{-k}}\right)^2
\end{aligned}
$$

with $\nu \in \{2+\log^{-1}(np_{-k}), 2+2\epsilon\}$. Also, since $\mathsf{E}(\sum_{\ell,m\in[p_{-k}]}\sum_{q\in[r_{-k}]} d_{\ell q}^{(k)} d_{mq}^{(k)} Z_{k,i\ell,t}Z_{k,jm,t}) = 0$ for $i \neq j$ due to Assumption 4 (i), we have

$$
\begin{aligned}
V_{2,2} &\lesssim \max_{i\in[p_k]} \frac{1}{p^2} \left(\frac{1}{n} \sum_{t\in[n]} \sum_{\ell\in[p_{-k}]} \sum_{q\in[r_{-k}]} d_{\ell q}^{(k)} d_{\ell q}^{(k)} \mathsf{E}(Z_{k,i\ell,t}^2)\right)^2 \\
&\lesssim \max_{i\in[p_k]} \frac{1}{p^2} \left(\frac{1}{n} \sum_{t\in[n]} (|\mathcal{F}_t|_2^2+\omega^2)\right)^2 \lesssim \frac{\omega^4}{p^2},
\end{aligned}
$$

due to Assumption 3 (ii) and Lemma B.10 (iv). Collecting the bounds on $V_{2,1}$ and $V_{2,2}$, we obtain by Markov's inequality,

$$
U_2 = O_P\left(\frac{\psi_{n,p}^{(k)}}{\sqrt{p}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{n}p_{-k}} \vee \frac{\omega^2}{p}\right),
$$

and thus

$$
U_1 + U_2 = O_P\left[\frac{\psi_{n,p}^{(k)}\bar{\psi}_{n,p}^{(k)}}{\sqrt{p_k}} \vee \frac{\bar{\psi}_{n,p}}{\sqrt{p}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\frac{1}{p_{-k}} \vee \bar{\psi}_{n,p}^{(k)}\right) \vee \frac{1}{p}\right]
$$

which completes the proof. $\square$

*Proof of Lemma B.12 (ii).* We continue with the decomposition in (B.18). WLOG, we fix $k=1$. Then from Lemma B.1 (ii), B.2 and B.3 and Assumption 5,

$$
\begin{aligned}
V_{1,1} \lesssim & \frac{1}{n^2p^2}\sum_{i,j\in[p_k]} \otimes_{i=2}^{K} \sum_{i_l,i_l'\in[p_l]}\sum_{t,u\in[n]} \|Z_{1,(i,\mathbf{i}_{2:K}),t}Z_{1,(j,\mathbf{i}'_{2:K}),t}\|_\nu \|Z_{1,(i,\mathbf{i}_{2:K}),u}Z_{1,(j,\mathbf{i}'_{2:K}),u}\|_\nu \\
& \times \exp\left(-\frac{c_0(\nu-2)(|t-u|+\sum_{l=2}^K |i_l-i_l'|)}{K\nu}\right) \\
\lesssim & \frac{(\tau_{n,p}^{(1)})^{\frac{4(\nu-1-\epsilon)}{\nu}}}{np_{-k}}\cdot\frac{1}{n}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^{\frac{2+2\epsilon}{\nu}}(|\mathcal{F}_u|_2+\omega)^{\frac{2+2\epsilon}{\nu}}\exp\left(-\frac{c_0|t-u|}{3K\log(np)}\right) \\
& \times \prod_{i=2}^{K}\frac{1}{p_l}\sum_{i_l,i_l'\in[p_l]}\exp\left(-\frac{c_0|i_l-i_l'|}{3K\log(np)}\right) \\
\lesssim & \frac{c_\epsilon K^K(\tau_{n,p}^{(1)})^{2-2\epsilon}}{np_{-k}}\log^K(np) \lesssim (\psi_{n,p}^{(1)})^2
\end{aligned}
$$

with $\nu = 2+\log^{-1}(np_{-k})$. Similarly, from Lemma B.1 (i) and Assumption 3 (ii),

$$
\begin{aligned}
V_{1,2} = & \frac{2}{n^2p^2}\otimes_{l\in[K]}\sum_{i_l,i_l'\in[p_l]}\left(\sum_{t\in[n]}\mathsf{E}(Z_{1,\mathbf{i},t}Z_{1,\mathbf{i}',t})\right)^2 \\
\lesssim & \frac{1}{n^2p^2}\otimes_{l\in[K]}\sum_{i_l,i_l'\in[p_l]}\left[\sum_{t\in[n]}\|Z_{1,\mathbf{i},t}\|_\nu\|Z_{1,\mathbf{i}'_{2:K},t}\|\exp\left(-\frac{c_0(\nu-2)\sum_{l\in[K]}|i_l-i_l'|}{K\nu}\right)\right]^2 \\
\lesssim & \frac{1}{p}\left[\frac{1}{n}\sum_{t\in[n]}(|\mathcal{F}_t|_2+\omega)^2\right]^2\prod_{l\in[K]}\frac{1}{p_l}\sum_{i_l,i_l'\in[p_l]}\exp\left(-\frac{2c_0\epsilon|i_l-i_l'|}{K(1+\epsilon)}\right)\lesssim\frac{C_\epsilon\omega^4}{p}.
\end{aligned}
$$

with $\nu=2+2\epsilon$. These arguments apply to all $k\in[K]$ and thus combining the bounds on $V_{1,1}$ and $V_{1,2}$ with (B.15) and (B.19), by Markov's inequality,

$$
U_1 = O_P\left[\left(\psi_{n,p}^{(k)}\vee\frac{1}{p}\right)\bar{\psi}_{n,p}^{(k)}\right].
$$

As for $U_2$, note that

$$
\begin{aligned}
\mathsf{E}(U_2^2) \le & 2\,\mathsf{E}\left(\left\|\frac{1}{np}\sum_{t\in[n]}\left(\mathbf{Z}_{1,t}\mathbf{D}_1\mathbf{D}_1^\top\mathbf{Z}_{1,t}^\top-\mathsf{E}\left(\mathbf{Z}_{1,t}\mathbf{D}_1\mathbf{D}_1^\top\mathbf{Z}_{1,t}^\top\right)\right)\right\|^2\right) \\
& +2\left\|\frac{1}{np}\sum_{t\in[n]}\mathsf{E}\left(\mathbf{Z}_{1,t}\mathbf{D}_1\mathbf{D}_1^\top\mathbf{Z}_{1,t}^\top\right)\right\|^2 =: 2\mathsf{E}(\|\mathbf{V}_1\|^2)+2\mathsf{E}(\|\mathbf{V}_2\|^2),
\end{aligned}
$$

with $\mathbf{V}_\ell = [V_{\ell,ij},\, i,j \in [p_k]]]$, $\ell = 1,2$. Setting $\nu = 2 + \log^{-1}(np)$, we have

$$\begin{aligned}
\|V_{1,ij}\|_2^2 \le& \frac{C^4 r_{-1}^2}{n^2 p^2 p_{-1}^2} \sum_{t,u\in[n]} \otimes_{l=2}^K \sum_{a_l,b_l,c_l,d_l\in[p_l]} \left|\mathsf{Cov}\left(Z_{1,(i,\mathbf{a}_{2:K}),t} Z_{1,(j,\mathbf{b}_{2:K}),t}, Z_{1,(i,\mathbf{c}_{2:K}),u} Z_{1,(j,\mathbf{d}_{2:K})u}\right)\right| \\
\le& \frac{8C^4 r_{-1}^2}{n^2 p^2 p_{-1}^2} \sum_{t,u\in[n]} \otimes_{l=2}^K \sum_{a_l,b_l,c_l,d_l\in[p_l]} \left\|Z_{1,(i,\mathbf{a}_{2:K}),t} Z_{1,(j,\mathbf{b}_{2:K}),t}\right\|_\nu \left\|Z_{1,(i,\mathbf{c}_{2:K}),u} Z_{1,(j,\mathbf{d}_{2:K})u}\right\|_\nu \times \\
& \exp\left\{-\frac{c_0(\nu-2)[|t-u| + \min(\sum_{l=2}^K |a_l - c_l|, \sum_{l=2}^K |b_l - c_l|, \sum_{l=2}^K |a_l - d_l|, \sum_{l=2}^K |b_l - d_l|)]}{K\nu}\right\} \\
\lesssim& \frac{(\tau_{n,p}^{(1)})^{\frac{4(\nu-1-\epsilon)}{\nu}}}{n^2 p^2 p_{-1}^2} \sum_{t,u\in[n]} \left(|\mathcal{F}_t|_2^{\frac{2+2\epsilon}{\nu}} + \omega^{\frac{2+2\epsilon}{\nu}}\right)\left(|\mathcal{F}_u|_2^{\frac{2+2\epsilon}{\nu}} + \omega^{\frac{2+2\epsilon}{\nu}}\right) \exp\left(-\frac{c_0|t-u|(\nu-2)}{K\nu}\right) \times \\
& \prod_{l=2}^K \sum_{a_l,b_l,c_l,d_l\in[p_l]} \exp\left(-\frac{c_0(\nu-2)\min(|a_l-c_l|, |b_l-c_l|, |a_l-d_l|, |b_l-d_l|)}{K\nu}\right) \\
\lesssim& \frac{(\tau_{n,p}^{(1)})^{2-2\epsilon}}{n^2 p^2 p_{-1}^2} \sum_{t,u\in[n]} (|\mathcal{F}_u|_2^{1+\epsilon} + \omega^{1+\epsilon})(|\mathcal{F}_u|_2^{1+\epsilon} + \omega^{1+\epsilon}) \exp\left(-\frac{c_0|t-u|)}{3K\log(np)}\right) \times \\
& \prod_{l=2}^K \sum_{a_l,b_l,c_l,d_l\in[p_l]} \exp\left(-\frac{c_0 \min(|a_l-c_l|, |b_l-c_l|, |a_l-d_l|, |b_l-d_l|)}{3K\log(np)}\right) \\
\lesssim& \frac{\omega^{2+2\epsilon}(\tau_{n,p}^{(1)})^{2-2\epsilon} p_{-1}\log^K(np)}{np^2} \lesssim \frac{(\psi_{n,p}^{(1)})^2}{p_1^2}
\end{aligned}$$

where the first inequality follows from Lemma B.10 (iv), the second from Assumption 5 (i) and Lemmas B.2 and B.3, the third from Lemma B.1 (ii) and the penultimate one from Assumption 5 (ii). Analogous arguments apply to all $k \in [K]$ such that

$$\mathsf{E}(\|\mathbf{V}_1\|^2) \le \sum_{i,j\in[p_k]} \|V_{1,ij}\|_2^2 \lesssim (\psi_{n,p}^{(k)})^2. \tag{B.20}$$

Next, continuing to fix $k = 1$, for all $i,j \in [p_1]$, it holds that

$$\begin{aligned}
V_{2,ij} \le& \frac{1}{np} \sum_{t\in[n]} \otimes_{l=2}^K \sum_{i_l,i'_l\in[p_l]} \sum_{q\in[r_{-1}]} d_{\ell q}^{(1)} d_{mq}^{(1)} \left|\mathsf{Cov}\left(Z_{1,(i,\mathbf{i}_{2:K}),t}, Z_{1,(j,\mathbf{i}'_{2:K}),t}\right)\right| \\
\le& \frac{C^2 r_{-1}}{npp_{-1}} \sum_{t\in[n]} \otimes_{l=2}^K \sum_{i_l,i'_l\in[p_l]} \|Z_{1,(j,\mathbf{i}_{2:K}),t}\|_\nu \|Z_{1,(i,\mathbf{i}'_{2:K}),t}\|_\nu \\
& \qquad \times \exp\left(-\frac{c_0(\nu-2)(|i-j| + \sum_{l=2}^K |i_l - i'_l|)}{K\nu}\right) \\
\le& \frac{C^2 r_{-1}}{npp_{-1}} \exp\left(-\frac{c_0\epsilon|i-j|}{K(1+\epsilon)\log(np)}\right) \sum_{t\in[n]} (|\mathcal{F}_t|_2^2 + \omega^2) \prod_{l=2}^K \sum_{i_l,i'_l\in[p_l]} \exp\left(-\frac{c_0\epsilon|i_l - i'_l|}{K(1+\epsilon)}\right)
\end{aligned}$$

$$\lesssim \frac{C_\epsilon \omega^2}{p} \exp\left(-\frac{c_0 \epsilon |i-j|}{K(1+\epsilon)}\right)$$

with $\nu = 2 + 2\epsilon$, where the second inequality follows from Lemma B.10 (iv) and Lemma B.3, the third from Lemma B.1 (i) and the last from Assumption 3 (ii). Similar arguments hold for all $k \in [K]$, and thus

$$\|\mathbf{V}_2\| \le \|\mathbf{V}_2\|_1 \le \sum_{j \in [p_k]} |V_{2,ij}| \lesssim \frac{C_\epsilon \omega^2}{p}. \tag{B.21}$$

Putting together the bounds on $\|\mathbf{V}_\ell\|$, $\ell = 1, 2$, we have

$$U_2 = O_P\left(\psi_{n,p}^{(k)} \vee \frac{1}{p}\right).$$

Combining the bound on $U_2$ with that on $U_1$, the proof is complete. □

**Lemma B.13.** *Let Assumptions 1, 2 and 3 hold. For each $k \in [K]$, we have the followings:*

*(i) Under Assumption 4,*

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| = O_P\left(M_n^{1-\epsilon}\left(\frac{\bar{\psi}_{n,p}^{(k)}}{\sqrt{n}} \vee \frac{1}{\sqrt{np_{-k}}}\right)\right),$$

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top \mathbf{Z}_{k,t}^\top\right\| = O_P\left(M_n^{1-\epsilon}\left(\frac{\bar{\psi}_{n,p}^{(k)}}{\sqrt{n}} \vee \frac{1}{\sqrt{np_{-k}}}\right)\right).$$

*(ii) Under Assumption 5,*

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| = O_P\left(M_n^{1-\epsilon}\left(\frac{\bar{\psi}_{n,p}^{(k)}}{\sqrt{n}} \vee \frac{1}{\sqrt{np_{-k}}}\right)\right),$$

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top \mathbf{Z}_{k,t}^\top\right\| = O_P\left(M_n^{1-\epsilon}\left(\frac{\bar{\psi}_{n,p}^{(k)}}{\sqrt{n}} \vee \frac{1}{\sqrt{np_{-k}}}\right)\right).$$

*Proof of Lemma B.13 (i).* Let us write

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| \le \frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\left(\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top - \mathbf{D}_k\mathbf{D}_k^\top\right)\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\|$$
$$+ \frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\mathbf{D}_k\mathbf{D}_k^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| =: U_1 + U_2. \tag{B.22}$$

By Lemma B.17, we have

$$U_1^2 \le \underbrace{\frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\left(\sum_{t\in[n]} Z_{k,i\ell,t}\mathsf{E}(X^{\mathsf{t}}_{k,jm,t})\right)^2}_{V_1}\left\|\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top - \mathbf{D}_k\mathbf{D}_k^\top\right\|^2$$

where, by Lemmas B.1 (i), B.2, and B.3 and Assumptions 3 (iii) and 4,

$$\begin{aligned}
\mathsf{E}(V_1) &= \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\sum_{t,u\in[n]}\mathsf{E}(X^{\mathsf{t}}_{k,jm,t})\mathsf{E}(X^{\mathsf{t}}_{k,jm,u})\mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,i\ell,u}) \\
&\lesssim \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\sum_{t,u\in[n]}\left|\mathsf{E}(X^{\mathsf{t}}_{k,jm,t})\mathsf{E}(X^{\mathsf{t}}_{k,jm,u})\right| \\
&\quad \times \|Z_{k,i\ell,t}\|_\nu\|Z_{k,i\ell,u}\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right) \\
&\lesssim \frac{1}{n^2}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^2(|\mathcal{F}_u|_2+\omega)^2\exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \le \frac{c_\epsilon M_n^{2-2\epsilon}\omega^{2+2\epsilon}}{n}
\end{aligned}$$

with $\nu = 2+2\epsilon$, which leads to

$$U_1 = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\bar{\psi}^{(k)}_{n,p}\right)$$

by Markov's inequality and (B.15). Similarly, additionally evoking Lemma B.10 (iv),

$$\begin{aligned}
\mathsf{E}(U_2^2) &\le \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\mathsf{E}\left[\left(\sum_{\ell,m\in[p_{-k}]}\sum_{q\in[r_{-k}]}\sum_{t\in[n]} d^{(k)}_{\ell q}d^{(k)}_{mq}\mathsf{E}(X^{\mathsf{t}}_{k,jm,t})Z_{k,i\ell,t}\right)^2\right] \\
&= \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m,m'\in[p_{-k}]}\sum_{q,q'\in[r_{-k}]}\sum_{t,u\in[n]} d^{(k)}_{\ell q}d^{(k)}_{mq}d^{(k)}_{\ell q'}d^{(k)}_{m'q'} \\
&\quad \times \mathsf{E}(X^{\mathsf{t}}_{k,jm,t})\mathsf{E}(X^{\mathsf{t}}_{k,jm',u})\mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,i\ell,u}) \\
&\lesssim \frac{1}{n^2p_{-k}}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^2(|\mathcal{F}_u|_2+\omega)^2\exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \lesssim \frac{c_\epsilon M_n^{2-2\epsilon}}{np_{-k}},
\end{aligned}$$

leading to

$$U_2 = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right),$$

which completes the proof of the first claim. The second claim follows analogously. □

*Proof of Lemma B.13 (ii).* We continue with the decomposition in (B.22) and fix $k = 1$. By

Lemmas B.1 (i), B.2, and B.3 and Assumptions 3 (iii) and 5,

$$
\begin{aligned}
\mathsf{E}(V_1) &= \frac{1}{n^2p^2}\sum_{i,j\in[p_1]}\otimes_{l=2}^K\sum_{i_l,i'_l\in[p_l]}\sum_{t,u\in[n]}\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{i}'_{2:K}),t})\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{i}'_{2:K}),u})\mathsf{Cov}(Z_{1,(i,\mathbf{i}_{2:K}),t},Z_{1,(i,\mathbf{i}_{2:K}),u})\\
&\lesssim \frac{1}{n^2p^2}\sum_{i,j\in[p_1]}\otimes_{l=2}^K\sum_{i_l,i'_l\in[p_l]}\sum_{t,u\in[n]}\left|\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{i}'_{2:K}),t})\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{i}'_{2:K}),u})\right|\\
&\qquad\times\|Z_{1,(i,\mathbf{i}_{2:K}),t}\|_\nu\|Z_{1,(i,\mathbf{i}_{2:K}),u}\|_\nu\exp\left(-\frac{c_0(\nu-2)|t-u|}{K\nu}\right)\\
&\lesssim \frac{1}{n^2}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^2(|\mathcal{F}_u|_2+\omega)^2\exp\left(-\frac{c_0\epsilon|t-u|}{K(1+\epsilon)}\right)\le\frac{c_\epsilon KM_n^{2-2\epsilon}\omega^{2+2\epsilon}}{n}
\end{aligned}
$$

with $\nu=2+2\epsilon$, which leads to

$$
U_1=O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\bar{\psi}^{(k)}_{n,p}\right)
$$

As for $U_2$, WLOG, we fix $k=1$ for notational convenience. Then,

$$
\begin{aligned}
\mathsf{E}(U_2^2) &\lesssim \frac{1}{n^2p^2p_{-1}^2}\sum_{i,j\in[p_1]}\otimes_{l=2}^K\sum_{a_l,b_l,c_l,d_l\in[p_l]}\sum_{t,u\in[n]}\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{c}_{2:K}),t})\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{d}_{2:K}),t})\\
&\qquad\times\mathsf{Cov}(Z_{1,(i,\mathbf{a}_{2:K}),t},Z_{1,(i,\mathbf{b}_{2:K}),u})\\
&\lesssim \frac{1}{n^2p^2p_{-1}^2}\sum_{i,j\in[p_1]}\otimes_{l=2}^K\sum_{a_l,b_l,c_l,d_l\in[p_l]}\sum_{t,u\in[n]}\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{c}_{2:K}),t})\mathsf{E}(X^{\mathsf{t}}_{1,(j,\mathbf{d}_{2:K}),t})\\
&\qquad\times\|Z_{1,(i,\mathbf{a}_{2:K}),t}\|_\nu\|Z_{1,(i,\mathbf{b}_{2:K}),u}\|_\nu\exp\left(-\frac{c_0(\nu-2)(|t-u|+\sum_{l=2}^K|i_l-i'_l|)}{K(\nu-2)}\right)\\
&\lesssim \frac{1}{np_{-1}}\cdot\frac{1}{n}\sum_{t,u\in[n]}(|\mathcal{F}_t|_2+\omega)^2(|\mathcal{F}_u|_2+\omega)^2\exp\left(-\frac{c_0\epsilon|t-u|}{K(1+\epsilon)}\right)\\
&\qquad\times\prod_{l=2}^K\frac{1}{p_l}\sum_{i_l,i'_l\in[p_l]}\exp\left(-\frac{c_0\epsilon|i_l-i'_l|}{K(1+\epsilon)}\right)\lesssim\frac{c_\epsilon K^K\omega^{2+2\epsilon}M_n^{2-2\epsilon}}{np_{-k}},
\end{aligned}
$$

with $\nu=2+2\epsilon$, by Assumptions 3 (iii) and 5 and Lemmas B.1 (i), B.3 and B.10 (iv). Hence by Markov's inequality,

$$
U_2=O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right).
$$

Combining the bound on $U_2$ with that on $U_1$, the proof of the first claim is complete. The second claim is proved analogously. □

**Lemma B.14.** *Let Assumptions 1, 2 and 3 hold. For each $k\in[K]$, we have the followings* *under either Assumptions 4 or 5:*

$$\frac{1}{np}\left\|\sum_{t\in[n]}\mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t}-\mathbf{X}_{k,t}\right)\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t}\right)^\top\right\| = O_P\left(\frac{M_n}{\tau^{(k)}_{n,p}}\psi^{(k)}_{n,p}\right),$$
$$\frac{1}{np}\left\|\sum_{t\in[n]}\mathsf{E}\left(\mathbf{X}_{k,t}\right)\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t}-\mathbf{X}_{k,t}\right)^\top\right\| = O_P\left(\frac{M_n}{\tau^{(k)}_{n,p}}\psi^{(k)}_{n,p}\right).$$

*Proof.* Let us write

$$\begin{aligned}&\frac{1}{np}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t}-\mathbf{X}_{k,t})\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top\mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t})^\top\right\|\\ &\le\frac{1}{np}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t}-\mathbf{X}_{k,t})\left(\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top-\mathbf{D}_k\mathbf{D}_k^\top\right)\mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t})^\top\right\|\\ &\quad+\frac{1}{np}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t}-\mathbf{X}_{k,t})\mathbf{D}_k\mathbf{D}_k^\top\mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t})^\top\right\| =: U_1+U_2.\end{aligned}$$

By Lemma B.17, we have

$$U_1^2\le\underbrace{\frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\left(\sum_{t\in[n]}\mathsf{E}(X^{\mathsf{t}}_{k,i\ell,t}-X_{k,i\ell,t})\mathsf{E}(X^{\mathsf{t}}_{k,jm,t})\right)^2}_{V_1}\left\|\widehat{\mathbf{D}}_k\widehat{\mathbf{D}}_k^\top-\mathbf{D}_k\mathbf{D}_k^\top\right\|^2$$

where, by Lemmas B.1 (i) and (iii) and Assumption 3 (ii) and (iii),

$$V_1\lesssim\left(\frac{1}{n(\tau^{(k)}_{n,p})^{1+2\epsilon}}\sum_{t\in[n]}(|\mathcal{F}_t|_2+\omega)^{3+2\epsilon}\right)^2\lesssim\left(\frac{M_n\omega^{2+2\epsilon}}{(\tau^{(k)}_{n,p})^{1+2\epsilon}}\right)^2=\left(\frac{M_n}{\tau^{(k)}_{n,p}}\psi^{(k)}_{n,p}\right)^2$$

which, together with (B.15), leads to

$$U_1=O_P\left(\frac{M_n}{\tau^{(k)}_{n,p}}\psi^{(k)}_{n,p}\bar{\psi}^{(k)}_{n,p}\right).$$

Similarly, additionally evoking Lemma B.10 (iv),

$$U_2^2\le\frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\left(\sum_{\ell,m\in[p_{-k}]}\sum_{q\in[r_{-k}]}\sum_{t\in[n]}d^{(k)}_{\ell q}d^{(k)}_{mq}\mathsf{E}(X^{\mathsf{t}}_{k,i\ell,t}-X_{k,i\ell,t})\mathsf{E}(X^{\mathsf{t}}_{k,jm,t})\right)^2$$

$$\lesssim \left( \frac{1}{n(\tau_{n,p}^{(k)})^{1+2\epsilon}} \sum_{t\in[n]} (|\mathcal{F}_t|_2 + \omega)^{3+2\epsilon} \right)^2 \lesssim \left( \frac{M_n}{\tau_{n,p}^{(k)}} \psi_{n,p}^{(k)} \right)^2,$$

leading to

$$U_2 = O\left( \frac{M_n}{\tau_{n,p}^{(k)}} \psi_{n,p}^{(k)} \right).$$

which completes the proof of the first claim. The second claim is proved analogously. □

**Lemma B.15.** *Let Assumptions 1, 2 and 3 hold. For each $k \in [K]$, we have the followings:*

*(i) Under Assumption 4,*

$$\left\| \frac{1}{np} \sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}) \left( \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right) \mathsf{E}(\mathbf{X}_{k,t})^\top \right\|$$
$$= O_P\left( \sum_{k'\in[K]\setminus\{k\}} \left( \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p_{k'}} \vee \frac{\psi_{n,p}^{(k')}}{\sqrt{p_{k'}}} \right) \right).$$

*(ii) Under Assumption 5,*

$$\left\| \frac{1}{np} \sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}) \left( \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right) \mathsf{E}(\mathbf{X}_{k,t})^\top \right\| = O_P\left( \sum_{k'\in[K]\setminus\{k\}} \left( \psi_{n,p}^{(k')} \vee \frac{1}{p_{k'}} \right) \right).$$

*Proof of Lemma B.15 (i).* By Lemma B.17, we have

$$\frac{1}{n^2p^2} \left\| \sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}) \left( \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right) \mathsf{E}(\mathbf{X}_{k,t})^\top \right\|_F^2$$
$$\leq \underbrace{\frac{1}{n^2p^2} \sum_{i,j\in[p_k]} \sum_{\ell,m\in[p_{-k}]} \left( \sum_{t\in[n]} \mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}) \mathsf{E}(X_{k,jm,t}^{\mathsf{t}}) \right)^2}_{U_1} \left\| \widehat{\mathbf{D}}_k \widehat{\mathbf{D}}_k^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right\|_F^2.$$

By Lemma B.1 (i) and Assumption 3 (ii), we have

$$U_1 \lesssim \left( \frac{1}{n} \sum_{t\in[n]} (|\mathcal{F}_t|_2 + \omega)^2 \right)^2 \lesssim \omega^4$$

which, in combination with (B.16), leads to the claim. □

*Proof of Lemma B.15 (ii).* The proof takes analogous steps as in that of Lemma B.15 (i) except that we evoke (B.17) in place of (B.16). □

**Lemma B.16.** *Let Assumptions 1, 2, 3, and 4 or 5 hold. For* $\breve{\mathbf{M}}_k^{[1]}(\tau) \in \mathbb{R}^{r_k \times r_k}$ *denotes the diagonal matrix containing the eigenvalues* $\breve{\mu}_j^{(k)}(\tau)$, $j \in [r_k]$, *of* $\breve{\mathbf{\Gamma}}^{(k),[1]}(\tau)$ *on its diagonal, we have*

$$\left\| \left( p_k^{-1} \breve{\mathbf{M}}_k^{[1]}(\tau) \right)^{-1} - \left( p_k^{-1} \mathbf{M}_{\chi,k} \right)^{-1} \right\| = O_P \left( \frac{1}{p_k} \left\| \breve{\mathbf{\Gamma}}^{(k),[1]} - \mathbf{\Gamma}_\chi^{(k)} \right\| \right).$$

*Proof.* Having $\mathbf{\Gamma}_\chi^{(k)}$ fulfil the condition (C1) in Lemma B.4 in place of $\mathbf{S}$ and Proposition B.2 take the role of (C2), the conclusions follow from Lemma B.4 (ii). □

**Lemma B.17.** *For some sequence of matrices* $\mathbf{A}_t = [a_{ij,t}], \mathbf{C}_t$, $t \in [n]$, *and some matrix* $\mathbf{B}$ *of compatible dimensions,*

$$\left\| \sum_{t \in [n]} \mathbf{A}_t \mathbf{B} \mathbf{C}_t \right\|_F^2 \le \sum_{i,j} \left\| \sum_{t \in [n]} a_{ij,t} \mathbf{C}_t \right\|_F^2 \|\mathbf{B}\|_F^2.$$

*Proof.* Writing $\mathbf{B} = [b_{j\ell}]$ and $\mathbf{C}_t = [c_{\ell m,t}]$, by Cauchy-Schwarz inequality,

$$\begin{aligned} \left\| \sum_{t \in [n]} \mathbf{ABC} \right\|_F^2 &= \sum_i \sum_m \left( \sum_j \sum_\ell \sum_{t \in [n]} a_{ij,t} b_{j\ell} c_{\ell m,t} \right)^2 \\ &\le \sum_i \sum_j \sum_\ell \sum_m \left( \sum_{t \in [n]} a_{ij,t} c_{\ell m,t} \right)^2 \cdot \|\mathbf{B}\|_F^2 = \sum_{i,j} \left\| \sum_{t \in [n]} a_{ij,t} \mathbf{C}_t \right\|_F^2 \|\mathbf{B}\|_F^2. \end{aligned}$$

□

## B.4 Proof of Theorem 3

Throughout, we suppress the dependence on $\tau$ where there is no confusion.

### B.4.1 Proof of Theorem 3 (i)

For each $k \in [K]$, analogously as in (B.13), we may write

$$\begin{aligned} \frac{1}{p_k} \breve{\mathbf{\Gamma}}^{(k),[2]} &= \frac{1}{np} \sum_{t \in [n]} \mathbf{Z}_{k,t} \breve{\mathbf{D}}_k^{[1]} (\breve{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \\ &+ \frac{1}{np} \sum_{t \in [n]} \mathbf{Z}_{k,t} \breve{\mathbf{D}}_k^{[1]} (\breve{\mathbf{D}}_k^{[1]})^\top \mathsf{E} \left( \mathbf{X}_{k,t}^{\mathsf{t}} \right)^\top + \frac{1}{np} \sum_{t \in [n]} \mathsf{E} \left( \mathbf{X}_{k,t}^{\mathsf{t}} \right) \breve{\mathbf{D}}_k^{[1]} (\breve{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{np} \sum_{t \in [n]} \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}\right) \breve{\mathbf{D}}^{[1]}_k (\breve{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t}\right)^\top \\
&+ \frac{1}{np} \sum_{t \in [n]} \mathsf{E}(\mathbf{X}_{k,t}) \breve{\mathbf{D}}^{[1]}_k (\breve{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}\right)^\top \\
&+ \frac{1}{np} \sum_{t \in [n]} \mathsf{E}(\mathbf{X}_{k,t}) \left( \breve{\mathbf{D}}^{[1]}_k (\breve{\mathbf{D}}^{[1]}_k)^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right) \mathsf{E}(\mathbf{X}_{k,t})^\top + \frac{1}{p_k} \boldsymbol{\Gamma}^{(k)}_\chi \\
&=: T_1 + T_{2,1} + T_{2,2} + T_{3,1} + T_{3,2} + T_4 + \frac{1}{p_k} \boldsymbol{\Gamma}^{(k)}_\chi, \qquad \text{(B.23)}
\end{aligned}
$$

where $\boldsymbol{\Gamma}^{(k)}_\chi$ is defined in (A.1). Then by (13), (B.31), (B.34), (B.39) and Lemma B.22, we have

$$
\frac{1}{p_k} \left\| \breve{\boldsymbol{\Gamma}}^{(k),[2]} - \boldsymbol{\Gamma}^{(k)}_\chi \right\| = O_P\left( \sum_{k' \in [K]} \frac{M_n^{1-\epsilon}}{\sqrt{n p_{-k'}}} \vee \frac{1}{p} \right). \qquad \text{(B.24)}
$$

Next, denote by $\breve{\widetilde{\mathbf{M}}}^{[2]}_k(\tau) \in \mathbb{R}^{r_k \times r_k}$ the diagonal matrix containing the eigenvalues $\breve{\mu}^{(k),[2]}_j(\tau)$, $j \in [r_k]$, of $\breve{\boldsymbol{\Gamma}}^{(k),[2]}(\tau)$ on its diagonal. By (B.24) and Lemma B.23 and the arguments analogous to those adopted in proving (B.4) and (B.5), we have $p_k^{-1} \breve{\widetilde{\mathbf{M}}}^{[2]}_k$ asymptotically invertible and

$$
\left\| \left( p_k^{-1} \breve{\widetilde{\mathbf{M}}}^{[2]}_k \right)^{-1} \right\| \le \frac{1}{\alpha^{(k)}_{r_k}} + o_P(1) = O_P(1). \qquad \text{(B.25)}
$$

Let us set

$$
\breve{\mathbf{H}}^{[2]}_k = \frac{1}{n\sqrt{p_k} p_{-k}} \sum_{t \in [n]} \mathsf{mat}_k(\mathcal{F}_t) \boldsymbol{\Delta}_k^\top \breve{\mathbf{D}}^{[1]}_k (\breve{\mathbf{D}}^{[1]}_k)^\top \boldsymbol{\Delta}_k \mathsf{mat}_k(\mathcal{F}_t)^\top \boldsymbol{\Lambda}_k^\top \breve{\mathbf{E}}^{[2]}_k \left( \frac{1}{p_k} \breve{\widetilde{\mathbf{M}}}^{[2]}_k \right)^{-1},
$$

such that by Assumption 1 (i),

$$
\frac{1}{\sqrt{p_k}} \boldsymbol{\Lambda}_k \breve{\mathbf{H}}^{[2]}_k = \frac{1}{np} \sum_{t \in [n]} \mathsf{E}(\mathbf{X}_{k,t}) \breve{\mathbf{D}}^{[1]}_k (\breve{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}(\mathbf{X}_{k,t})^\top \breve{\mathbf{E}}^{[2]}_k \left( \frac{1}{p_k} \breve{\widetilde{\mathbf{M}}}^{[2]}_k \right)^{-1}.
$$

Then from (B.23), we may write

$$
\breve{\mathbf{E}}^{[2]}_k - \frac{1}{\sqrt{p_k}} \boldsymbol{\Lambda}_k \breve{\mathbf{H}}^{[2]}_k = \left( T_1 + T_{2,1} + T_{2,2} + T_{3,1} + T_{3,2} \right) \breve{\mathbf{E}}^{[2]}_k \left( \frac{1}{p_k} \breve{\widetilde{\mathbf{M}}}^{[2]}_k \right)^{-1},
$$

from which we derive that

$$
\left\| \breve{\mathbf{E}}^{[2]}_k - \frac{1}{\sqrt{p_k}} \boldsymbol{\Lambda}_k \breve{\mathbf{H}}^{[2]}_k \right\| = O_P\left[ \frac{M_n^{1-\epsilon}}{\sqrt{n p_{-k}}} \vee \frac{1}{p} \vee \frac{\psi^{(k)}_{n,p}}{\sqrt{p}} \vee \sum_{k' \in [K] \setminus \{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{n p_{-k'}}} \left( \frac{M_n^{1-\epsilon}}{\sqrt{n}} + \frac{\psi^{(k)}}{\sqrt{p_k}} \right) \right]
$$

$$= O_P\left(\sum_{k'\in[K]} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right)$$

by (B.25), (B.31), (B.34), (B.39) and (13). Finally, the conclusion follows from that

$$\|\breve{\mathbf{H}}_k^{[2]}\| \leq \frac{1}{n\sqrt{p_k}p_{-k}} \sum_{t\in[n]} |\mathcal{F}_t|_2^2 \|\boldsymbol{\Delta}_k\|^2 \|\boldsymbol{\Lambda}_k\| \left\| \left(\frac{1}{p_k}\widetilde{\breve{\mathbf{M}}}_k^{[2]}\right)^{-1} \right\| \leq (\alpha_{r_k}^{(k)})^{-1}\omega^2(1+o_P(1)) \quad \text{(B.26)}$$

by Assumptions 1 and 3 (ii) and (B.25), and hence

$$\mathbf{I}_{r_k} = (\breve{\mathbf{E}}_k^{[2]})^\top \breve{\mathbf{E}}_k^{[2]} = \frac{1}{p_k}(\breve{\mathbf{H}}_k^{[2]})^\top \boldsymbol{\Lambda}_k^\top \boldsymbol{\Lambda}_k \breve{\mathbf{H}}_k^{[2]} + o_P(1) = (\breve{\mathbf{H}}_k^{[2]})^\top \breve{\mathbf{H}}_k^{[2]} + o_P(1).$$

### B.4.2 Proof of Theorem 3 (ii)

For any $k \in [K]$ and $i \in [p_k]$, we have

$$\begin{aligned}
&\sqrt{np_{-k}}\left(\breve{\boldsymbol{\Lambda}}_{k,i\cdot}^{[2]} - \boldsymbol{\Lambda}_{k,i\cdot}\breve{\mathbf{H}}_k^{[2]}\right) \\
&= \frac{1}{\sqrt{np}} \sum_{t\in[n]} \Big[\mathbf{Z}_{k,i\cdot t}\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top + \mathbf{Z}_{k,i\cdot,t}\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top \\
&\quad + \mathsf{E}\left(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}}\right)\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top + \mathsf{E}\left(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}} - \mathbf{X}_{k,i\cdot,t}\right)\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top \\
&\quad + \mathsf{E}(\mathbf{X}_{k,i\cdot,t})\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t}\right)^\top\Big] \left(\breve{\mathbf{E}}_k^{[2]}\left(\frac{1}{p_k}\widetilde{\breve{\mathbf{M}}}_k^{[2]}\right)^{-1} - \mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right) \\
&+ \frac{1}{\sqrt{np}} \sum_{t\in[n]} \Big[\mathbf{Z}_{k,i\cdot t}\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top + \mathsf{E}\left(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}}\right)\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \\
&\quad + \mathsf{E}\left(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}} - \mathbf{X}_{k,i\cdot,t}\right)\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top \\
&\quad + \mathsf{E}(\mathbf{X}_{k,i\cdot,t})\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t}\right)^\top\Big] \mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1} \\
&+ \frac{1}{\sqrt{np}} \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t}\left(\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right)\mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}}\right)^\top \mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1} \\
&+ \frac{1}{\sqrt{np}p_{-k}} \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top \mathsf{E}\left(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t}\right)^\top \mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1} \\
&+ \frac{1}{\sqrt{np}p_{-k}} \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top \mathsf{E}\left(\mathbf{X}_{k,t}\right)^\top \mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1} \\
&=: U_1 + U_2 + U_3 + U_4 + \frac{1}{\sqrt{np_{-k}}} \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t}\boldsymbol{\Delta}_k \mathsf{mat}_k(\mathcal{F}_t)^\top (\boldsymbol{\Lambda}_k^\circ)^\top \breve{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}, \qquad \text{(B.27)}
\end{aligned}$$

where $\mathbf{E}_{\chi,k}$ and $\breve{\mathbf{J}}_k^{[2]}$ are defined in Lemma B.24, and $\mathbf{\Lambda}_k^\circ = p_k^{-1/2}\mathbf{E}_{\chi,k}^\top \mathbf{\Lambda}_k \in \mathbb{R}^{r_k \times r_k}$. Under the conditions made in (13), by (B.32), (B.35), (B.40) and Lemma B.24 (iii), we have

$$\frac{U_1}{\sqrt{np_{-k}}} = O_P\left[\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p}\right)\left(\sum_{k' \in [K]} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right)\right].$$

Similarly, (B.33), (B.36), and (B.41) give

$$\frac{U_2}{\sqrt{np_{-k}}} = O_P\left[\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\left(\frac{1}{\sqrt{p_k}} \vee \frac{M_n^\epsilon \sqrt{\log(np_{-k})}}{(\tau_{n,p}^{(k)})^\epsilon}\right) \vee \frac{1}{p}\right],$$

(B.37) leads to

$$\frac{U_3}{\sqrt{np_{-k}}} = O_P\left[\frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\sum_{k' \in [K]\setminus\{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right)\right]$$

and finally, by (B.38),

$$\frac{U_4}{\sqrt{np_{-k}}} = O_P\left[\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\left(\frac{M_n}{\tau_{n,p}^{(k)}}\right)^{1+2\epsilon}\right].$$

Then under the additional conditions made in Theorem 3 (ii), namely that $\sqrt{np_{-k}} = o(p)$ and $M_n = M$, we have

$$\max(U_1, U_2, U_3, U_4) = o_P(1) \quad \text{as} \quad \min(n, p_1, \ldots, p_K) \to \infty. \tag{B.28}$$

Next, let us define $\mathbf{Y}_t \in \mathbb{R}^{r_k}$ as

$$\mathbf{Y}_t = \left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1} \breve{\mathbf{J}}_k^{[2]} \mathbf{\Lambda}_k^\circ \mathsf{mat}_k(\mathcal{F}_t)\left(\frac{1}{\sqrt{p_{-k}}}\mathbf{\Delta}_k\right)^\top \mathbf{Z}_{k,i\cdot,t}^\top.$$

In what follows, we derive the asymptotic distribution of $n^{-1/2}\sum_{t\in[n]}\mathbf{Y}_t$ which, in view of the decomposition in (B.27), determines the asymptotic distribution of $\breve{\mathbf{\Lambda}}_{k,i\cdot}^{[2]}$ after appropriate centering and scaling.

Recalling that $\mathbf{\Psi}_{i,tu}^{(k)} = \mathrm{diag}(\mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,i\ell,u}),\, \ell \in [p_{-k}])$, we have

$$\left\|\mathbf{\Psi}_{i,tu}^{(k)}\right\| \le \max_{\ell \in [p_{-k}]} |\mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,i\ell,u})| \lesssim (|\mathcal{F}_t|_2 + \omega)(|\mathcal{F}_u|_2 + \omega)\exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right),$$

by Lemmas B.1 (i), B.2 and B.3. Then,

$$\mathbf{\Phi}_i^{(k)} := \mathsf{Cov}\left(\frac{1}{\sqrt{n}}\sum_{t\in[n]}\mathbf{Y}_t\right)$$
$$= \frac{1}{np}\sum_{t,u\in[n]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\breve{\mathbf{J}}_k^{[2]}\mathbf{E}_{\chi,k}^\top\mathbf{\Lambda}_k\mathsf{mat}_k(\mathcal{F}_t)\mathbf{\Delta}_k^\top\mathbf{\Psi}_{i,tu}^{(k)}\mathbf{\Delta}_k\mathsf{mat}_k(\mathcal{F}_u)^\top\mathbf{\Lambda}_k^\top\mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}$$
$$= \frac{1}{p_k}(\mathbf{\Gamma}_f^{(k)})^{-1}\mathbf{E}_{\chi,k}\mathbf{\Lambda}_k\left(\frac{1}{np_{-k}}\sum_{t,u\in[n]}\mathsf{mat}_k(\mathcal{F}_t)\mathbf{\Delta}_k^\top\mathbf{\Psi}_{i,tu}^{(k)}\mathbf{\Delta}_k\mathsf{mat}_k(\mathcal{F}_u)^\top\right)\mathbf{\Lambda}_k^\top\mathbf{E}_{\chi,k}(\mathbf{\Gamma}_f^{(k)})^{-1} + o(1)$$
$$= (\mathbf{\Gamma}_f^{(k)})^{-1}\left(\frac{1}{np_{-k}}\sum_{t,u\in[n]}\mathsf{mat}_k(\mathcal{F}_t)\mathbf{\Delta}_k^\top\mathbf{\Psi}_{i,tu}^{(k)}\mathbf{\Delta}_k\mathsf{mat}_k(\mathcal{F}_u)^\top\right)(\mathbf{\Gamma}_f^{(k)})^{-1} + o(1)$$

by Assumptions 1 (i), 2 and 4 (ii), and Lemmas A.1 and B.24 (iv). Also, it can be shown that

$$\left\|\mathbf{\Phi}_i^{(k)}\right\| \le \left\|\mathbf{\Gamma}_f^{(k)}\right\|^{-2}\cdot\frac{1}{n}\sum_{t,u\in[n]}|\mathcal{F}_t|_2|\mathcal{F}_u|_2(|\mathcal{F}_t|_2+\omega)(|\mathcal{F}_u|_2+\omega)\exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) + o(1)$$
$$\lesssim M^{2-2\epsilon}\omega^{2+2\epsilon}c_\epsilon,$$

from Assumptions 2 and 3 (iii). We now verify the conditions in Equations (3), (7) and (8) of Merlevede and Peligrad (2020) for $\mathbf{a}^\top\mathbf{Y}_t$ with any $\mathbf{a}\in\mathbb{B}_2(1)$, where $\mathbb{B}_2(1) := \{\mathbf{b}\in\mathbb{R}^{r_k} : |\mathbf{b}|_2 = 1\}$.

**Equation (3).** By Assumptions 1, 3 (ii) and (iii) and 4 (i) and Lemmas A.1, B.1 (i) and B.24 (iv), for any $j\in[r_k]$ and $\mathbf{a}\in\mathbb{B}_2(1)$,

$$\frac{1}{n}\sum_{t\in[n]}\|\mathbf{a}^\top\mathbf{Y}_t\|_2^2 \le \frac{1}{np}\left\|\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right\|^2\left\|\mathbf{E}_{\chi,k}^\top\mathbf{\Lambda}_k\right\|^2\cdot\sum_{t\in[n]}|\mathcal{F}_t|_2^2\left\|\mathbf{\Delta}_k^\top\mathbf{Z}_{k,i\cdot,t}^\top\right\|_2^2$$
$$\lesssim \frac{1}{np_{-k}}\sum_{t\in[n]}|\mathcal{F}_t|_2^2\sum_{\ell\in[p_{-k}]}\sum_{q\in[r_{-k}]}\delta_{k,\ell q}^2\|Z_{k,i\ell,t}\|_2^2$$
$$\lesssim \frac{1}{n}\sum_{t\in[n]}|\mathcal{F}_t|^2(|\mathcal{F}_t|_2+\omega)^2 \lesssim \omega^{2+2\epsilon}M^{2-2\epsilon}. \tag{B.29}$$

By using the arguments analogous to those leading to (B.29), we derive that

$$\|\mathbf{a}^\top\mathbf{Y}_t\|_4^4 \le \left\|\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right\|^4\left\|\frac{1}{\sqrt{p_k}}\mathbf{E}_{\chi,k}^\top\mathbf{\Lambda}_k\right\|^4|\mathcal{F}_t|_2^4\mathsf{E}\left\{\frac{1}{p_{-k}^2}\left[\sum_{q\in[r_{-k}]}\left(\sum_{\ell\in[p_{-k}]}\delta_{\chi,\ell q}Z_{k,i\ell,t}\right)^2\right]^2\right\}$$

$$\lesssim \frac{|\mathcal{F}_t|_2^4}{p_{-k}^2} \sum_{q,q' \in [r_{-k}]} \sum_{\substack{\ell,\ell',\ell'',\ell''' \in [p_{-k}] \\ \{\ell,\ell'\}=\{\ell'',\ell'''\} \text{ or} \\ \{\ell,\ell''\}=\{\ell',\ell'''\} \text{ or} \\ \{\ell,\ell'''\}=\{\ell',\ell''\}}} \delta_{k,\ell q}\delta_{k,\ell' q}\delta_{k,\ell'' q'}\delta_{k,\ell''' q'} \mathsf{E}(Z_{k,i\ell,t}Z_{k,i\ell',t}Z_{k,i\ell'',t}Z_{k,i\ell''',t})$$

$$\lesssim |\mathcal{F}_t|_2^4(|\mathcal{F}_t|_2 + \omega)^4. \tag{B.30}$$

Also for any $\varepsilon > 0$,

$$\mathsf{P}\left(|\mathbf{a}^\top \mathbf{Y}_t| > \sqrt{n}\varepsilon\right) \le \frac{\|\mathbf{a}^\top \mathbf{Y}_t\|_2^2}{n\varepsilon^2} \lesssim \frac{|\mathcal{F}_t|_2^2(|\mathcal{F}_t|_2 + \omega)^2}{n\varepsilon^2}.$$

Combined, we derive that

$$\frac{1}{n}\sum_{t\in[n]} \mathsf{E}\left(|\mathbf{a}^\top \mathbf{Y}_t|^2 \cdot \mathbb{I}_{\{|\mathbf{a}^\top \mathbf{Y}_t| > \sqrt{n}\varepsilon\}}\right) \le \frac{1}{n}\sum_{t\in[n]} \sqrt{\|\mathbf{a}^\top \mathbf{Y}_t\|_4^4 \cdot \mathsf{P}(|\mathbf{a}^\top \mathbf{Y}_t| > \sqrt{n}\varepsilon)}$$
$$\lesssim \frac{1}{n}\sum_{t\in[n]} \frac{|\mathcal{F}_t|_2^3(|\mathcal{F}_t|_2 + \omega)^3}{\sqrt{n}\varepsilon} \lesssim \frac{M^{4-2\epsilon}\omega^{2+2\epsilon}}{\sqrt{n}\varepsilon} \to 0$$

as $n \to \infty$, where the first inequality follows from Hölder's inequality.

**Equations (7) and (8).** Using the arguments adopted in (B.30), for any $j \in [r_k]$ with $\delta = 2$,

$$\frac{1}{n}\sum_{t\in[n]} \|\mathbf{a}^\top \mathbf{Y}_t\|_{2+\delta}^2 \lesssim \frac{1}{n}\sum_{t\in[n]} |\mathcal{F}_t|_2^2(|\mathcal{F}_t|_2 + \omega)^2 \lesssim \omega^{2+2\epsilon}M^{2-2\epsilon}$$

for all $n \ge 1$. The second claim in Equation (7) and that in (8) are met by Assumption 4 (i) and Lemma B.2, due to the discussions in Section 2.1.1 of Merlevede and Peligrad (2020). Equipped with the above, we can show that by Corollary 2.2 of Merlevede and Peligrad (2020), we have for any $\mathbf{a} \in \mathbb{B}_2(1)$,

$$\frac{1}{\sqrt{n}}\sum_{t\in[n]} \mathbf{a}^\top \mathbf{Y}_t \to \mathcal{N}_{r_k}\left(\mathbf{0}, \mathbf{a}^\top \mathbf{\Phi}_i^{(k)} \mathbf{a}\right)$$

as $\min(n, p_1, \dots, p_K) \to \infty$. This, combined with Cramér-Wold theorem (cf. Theorem 29.4 of Billingsley, 1995), completes the proof in combination with (B.28).

### B.4.3 Supporting lemmas

**Lemma B.18.** *Let the conditions in Theorem 3 hold. Then for all $k \in [K]$,*

$$\left\| \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \mathbf{\Delta}_k \mathbf{\Delta}_k^\top \right\| = O_P \left( \sum_{k' \in [K] \setminus \{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p} \right).$$

*Proof.* The proof proceeds analogously as in the proof of Lemmas B.10 (iii) and B.11 with Theorem 2 (and (13)) in place of Theorem 1. □

In proving Lemmas B.19–B.21, we omit that the results are derived under the conditions made in Theorem 3 and also the dependence on $\tau$.

**Lemma B.19.** *For each $k \in [K]$, we have*

$$\frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \right\| \\ = O_P \left[ \left( \frac{M_n^{1-\epsilon}}{\sqrt{n}} \vee \frac{1}{\sqrt{p}} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \right) \sum_{k' \in [K] \setminus \{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{n} p_{-k}} \vee \frac{1}{p} \right]. \tag{B.31}$$

*Also, for any $i \in [p_k]$,*

$$\frac{\sqrt{p_k}}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,i\cdot,t} \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \right\| \\ = O_P \left[ \left( \frac{M_n^{1-\epsilon}}{\sqrt{n}} \vee \frac{1}{\sqrt{p}} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \right) \sum_{k' \in [K] \setminus \{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{n} p_{-k}} \vee \frac{1}{p} \right], \tag{B.32}$$

$$\frac{\sqrt{p_k}}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,i\cdot,t} \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \mathbf{E}_{\chi,k} \check{\mathbf{J}}_k^{[2]} \left( \frac{1}{p_k} \mathbf{M}_{\chi,k} \right)^{-1} \right\| \\ = O_P \left[ \frac{1}{\sqrt{p_k}} \left( \frac{M_n^{1-\epsilon}}{\sqrt{n}} \vee \frac{1}{\sqrt{p}} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \right) \sum_{k' \in [K] \setminus \{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} + \frac{1}{\sqrt{p}} \left( \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \right) \vee \frac{1}{p} \right]. \tag{B.33}$$

*Proof.* To prove (B.31), let us write

$$\frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \right\| \le \frac{1}{np} \left\| \sum_{t \in [n]} \mathbf{Z}_{k,t} \left( \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \mathbf{\Delta}_k \mathbf{\Delta}_k^\top \right) \mathbf{Z}_{k,t}^\top \right\| \\ + \frac{1}{np} \left\| \frac{1}{p_{-k}} \sum_{t \in [n]} \mathbf{Z}_{k,t} \mathbf{\Delta}_k \mathbf{\Delta}_k^\top \mathbf{Z}_{k,t}^\top \right\| =: U_1 + U_2.$$

By Lemma B.17,

$$U_1^2 \leq \underbrace{\frac{1}{n^2p^2} \sum_{i,j\in[p_k]} \sum_{\ell,m\in[p_{-k}]} \left( \sum_{t\in[n]} Z_{k,i\ell,t} Z_{k,jm,t} \right)^2}_{V_1} \left\| \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right\|^2,$$

where $V_1$ is bounded analogously as the corresponding term in the proof of Lemma B.12 (i). Hence, with Lemma B.18, by Markov's inequality,

$$U_1 = O_P \left[ \left( \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{n}} \vee \frac{\omega^2}{\sqrt{p}} \right) \left( \sum_{k'\in[K]\backslash\{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p} \right) \right].$$

As for $U_2$, we adopt the analogous arguments as those employed in bounding $U_2$ in the proof of Lemma B.12 (i) in combination with Lemma B.10 (iv), which yields

$$U_2 = O_P \left( \frac{\psi_{n,p}^{(k)}}{\sqrt{p}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{\omega^2}{p} \right),$$

and thus

$$U_1 + U_2 = O_P \left[ \left( \frac{M_n^{1-\epsilon}}{\sqrt{n}} \vee \frac{1}{\sqrt{p}} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \right) \sum_{k'\in[K]\backslash\{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{\psi_{n,p}^{(k)}}{\sqrt{p}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p} \right],$$

which completes the proof of the first claim. The claim in (B.32) follows following the analogous steps.

For (B.33), we proceed similarly with some modifications. Let us write

$$\begin{aligned}
&\frac{\sqrt{p_k}}{np} \left\| \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t} \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \mathbf{E}_{\chi,k} \right\| \\
&\leq \frac{\sqrt{p_k}}{np} \left\| \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t} \left( \check{\mathbf{D}}_k^{[1]} (\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \right) \mathbf{Z}_{k,t}^\top \boldsymbol{\Lambda}_k \right\| \\
&\qquad + \frac{\sqrt{p_k}}{npp_{-k}} \left\| \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top \mathbf{Z}_{k,t}^\top \mathbf{E}_{\chi,k} \right\| =: U_3 + U_4.
\end{aligned}$$

By Lemma B.17, with $\mathbf{E}_{\chi,k} = [e_{\chi,ij}^{(k)}, i \in [p_k], j \in [r_k]]$,

$$U_3^2 \leq \underbrace{\frac{1}{n^2 p_k p_{-k}^2} \sum_{q \in [r_k]} \sum_{\ell, m \in [p_{-k}]} \left( \sum_{j \in [p_k]} \sum_{t \in [n]} Z_{k,i\ell,t} Z_{k,jm,t} e_{\chi,jq}^{(k)} \right)^2}_{V_3} \left\| \breve{\mathbf{D}}_k^{[1]} (\breve{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \mathbf{\Delta}_k \mathbf{\Delta}_k^\top \right\|^2,$$

where

$$\begin{aligned} \mathsf{E}(V_3) \leq & \frac{2}{n^2 p_k p_{-k}^2} \sum_{q \in [r_k]} \sum_{\ell, m \in [p_{-k}]} \sum_{j, j' \in [p_k]} \sum_{t, u \in [n]} e_{\chi,jq}^{(k)} e_{\chi,j'q}^{(k)} \mathsf{Cov}(Z_{k,i\ell,t} Z_{k,jm,t}, Z_{k,i\ell,u} Z_{k,j'm,u}) \\ & + \frac{2}{n^2 p_k p_{-k}^2} \sum_{q \in [r_k]} \sum_{\ell, m \in [p_{-k}]} \left( \sum_{j \in [p_k]} \sum_{t \in [n]} e_{\chi,jq}^{(k)} \mathsf{E}(Z_{k,i\ell,t} Z_{k,jm,t}) \right)^2 =: V_{3,1} + V_{3,2}. \end{aligned}$$

From Lemmas B.1 (i) and (ii), B.2, B.3 and B.24 (i) and Assumptions 3 (iii) and 4,

$$\begin{aligned} V_{3,1} \lesssim & \frac{1}{n^2 p^2} \sum_{\ell \in [p_{-k}]} \sum_{t, u \in [n]} \|Z_{k,i\ell,t}^2\|_\nu \|Z_{k,i\ell,u}^2\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right) \\ & + \frac{1}{n^2 p^2} \sum_{j \in [p_k]} \sum_{\ell, m \in [p_{-k}]} \sum_{t, u \in [n]} \|Z_{k,i\ell,t}\|_\nu \|Z_{k,jm,t}\|_\nu \|Z_{k,i\ell,u}\|_\nu \|Z_{k,jm,u}\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right) \\ \lesssim & \frac{(\tau_{n,p}^{(k)})^{2-2\epsilon}}{n^2 p_k^2 p_{-k}} \sum_{t, u \in [n]} (|\mathcal{F}_t|_2 + \omega)^{1+\epsilon} (|\mathcal{F}_u|_2 + \omega)^{1+\epsilon} \exp\left(-\frac{c_0|t-u|}{3\log(np_{-k})}\right) \\ & + \frac{1}{n^2 p_k} \sum_{t, u \in [n]} (|\mathcal{F}_t|_2 + \omega)^2 (|\mathcal{F}_u|_2 + \omega)^2 \exp\left(-\frac{c_0 \epsilon |t-u|}{1+\epsilon}\right) \\ \lesssim & \frac{c_\epsilon (\tau_{n,p}^{(k)})^{2-2\epsilon}}{p_k^2} \cdot \frac{\log(np_{-k})}{np_{-k}} + \frac{M_n^{2-2\epsilon} c_\epsilon}{np_k} \lesssim \left(\frac{\psi_{n,p}^{(k)}}{p_k} + \frac{M_n^{1-\epsilon}}{\sqrt{np_k}}\right)^2 \end{aligned}$$

with $\nu \in \{2 + \log^{-1}(np_{-k}), 2 + 2\epsilon\}$ for the case of $j = i$ and $j \neq i$, respectively. Also by Lemma B.1 (i) and Assumption 4, we have

$$V_{3,2} \lesssim \frac{1}{n^2 p^2} \sum_{\ell \in [p_{-k}]} \left( \sum_{t \in [n]} \mathsf{E}(Z_{k,i\ell,t}^2) \right)^2 \lesssim \frac{1}{p_k p} \left( \frac{1}{n} \sum_{t \in [n]} (|\mathcal{F}_t|_2^2 + \omega^2) \right)^2 \lesssim \frac{\omega^4}{p_k p}.$$

Combining the bounds on $V_{3,1}$ and $V_{3,2}$ with Lemma B.18, we have by Markov's inequality,

$$U_3 = O_P \left[ \frac{1}{\sqrt{p_k}} \left( \frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{n}} \vee \frac{\omega^2}{\sqrt{p}} \right) \left( \sum_{k' \in [K] \setminus \{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p} \right) \right].$$

As for $U_4$, writing $\boldsymbol{\Delta}_k = [\delta_{k,\ell q},\ \ell \in [p_{-k}],\ q \in [r_{-k}]]$, we can upper bound $\mathsf{E}(U_4^2)$ by

$$\frac{2}{n^2 p_k p_{-k}^4} \sum_{q\in[r_k]} \mathsf{E}\left[\left(\sum_{t\in[n]}\sum_{j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\sum_{q'\in[r_{-k}]} e_{\chi,jq}^{(k)}\delta_{k,\ell q'}\delta_{k,mq'}\left(Z_{k,i\ell,t}Z_{k,jm,t} - \mathsf{E}(Z_{k,i\ell,t}Z_{k,jm,t})\right)\right)^2\right]$$
$$+ \frac{2}{n^2 p_k p_{-k}^4}\left\|\sum_{t\in[n]} \mathsf{E}\left(\mathbf{Z}_{k,t}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top \mathbf{Z}_{k,t}^\top \mathbf{E}_{\chi,k}\right)\right\|^2 =: V_{4,1} + V_{4,2}.$$

Then, by Assumptions 3 (iii) and 4, Lemmas B.1 (i) and (ii), B.2, B.3, B.10 (iv) and B.24 (i),

$$\begin{aligned}
V_{4,1} &\lesssim \frac{1}{n^2 p_k^2 p_{-k}^4} \sum_{j\in[p_k]}\sum_{\ell,\ell',m,m'\in[p_{-k}]}\sum_{t,u\in[n]} \mathsf{Cov}(Z_{k,i\ell,t}Z_{k,jm,t}, Z_{k,i\ell',u}Z_{k,jm',u}) \\
&\le \frac{1}{n^2 p_k^2 p_{-k}^4} \sum_{\substack{\ell,\ell',m,m'\in[p_{-k}] \\ (\ell,m)=(\ell',m') \text{ or} \\ (\ell,m)=(m',\ell') \text{ or} \\ (\ell,\ell')=(m,m')}} \sum_{t,u\in[n]} \|Z_{k,i\ell,t}Z_{k,im,t}\|_\nu \|Z_{k,i\ell',u}Z_{k,im',u}\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right) \\
&\quad + \frac{1}{n^2 p_k^2 p_{-k}^4} \sum_{\substack{j\in[p_k] \\ j\ne i}}\sum_{\ell,m\in[p_{-k}]}\sum_{t,u\in[n]} \|Z_{k,i\ell,t}\|_\nu \|Z_{k,jm,t}\|_\nu \|Z_{k,i\ell,u}\|_\nu \|Z_{k,jm,u}\|_\nu \exp\left(-\frac{c_0(\nu-2)|t-u|}{\nu}\right) \\
&\lesssim \frac{(\tau_{n,p}^{(k)})^{2-2\epsilon}}{n^2 p_k^2 p_{-k}^2} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^{1+\epsilon}(|\mathcal{F}_u|_2+\omega)^{1+\epsilon} \exp\left(-\frac{c_0|t-u|}{3\log(np_{-k})}\right) \\
&\quad + \frac{M_n^{2-2\epsilon}}{n^2 p_k p_{-k}^2} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^{1+\epsilon}(|\mathcal{F}_u|_2+\omega)^{1+\epsilon} \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \\
&\lesssim \left(\frac{\psi_{n,p}^{(k)}}{\sqrt{p_k p}} + \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}p}}\right)^2
\end{aligned}$$

with $\nu \in \{2+\log^{-1}(np_{-k}), 2+2\epsilon\}$ for when $i = j$ and $i \ne j$, respectively. Besides, we have

$$V_{4,2} \lesssim \frac{1}{p^2}\left[\frac{1}{np_{-k}}\sum_{t\in[n]}\sum_{\ell\in[p_{-k}]} \mathsf{E}(Z_{k,i\ell,t}^2)\right]^2 \lesssim \frac{1}{p^2}\left(\frac{1}{n}\sum_{t\in[n]}(|\mathcal{F}_t|_2^2+\omega^2)\right)^2 \lesssim \frac{\omega^4}{p^2},$$

due to Assumptions 1, 3 (ii), 4 (i) and Lemma B.10 (iv). Collecting the bounds on $V_{4,1}$ and $V_{4,2}$, we obtain by Markov's inequality,

$$U_4 = O_P\left[\frac{1}{\sqrt{p}}\left(\frac{\psi_{n,p}^{(k)}}{\sqrt{p_k}} \vee \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right) \vee \frac{\omega^2}{p}\right].$$

Combining the bounds on $U_3$ and $U_4$ and evoking Lemma A.1 and (B.26) completes the proof. □

**Lemma B.20.** *For each $k \in [K]$, we have*

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{\sqrt{p_{-k}}}\right)\right), \quad \text{(B.34a)}$$

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{\sqrt{p_{-k}}}\right)\right). \quad \text{(B.34b)}$$

*Also, for any $i \in [p_k]$,*

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t}\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{\sqrt{p_{-k}}}\right)\right), \quad \text{(B.35a)}$$

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}})\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{\sqrt{p_{-k}}}\right)\right), \quad \text{(B.35b)}$$

*and*

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}})\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top \mathbf{Z}_{k,t}^\top \mathbf{E}_{\chi,k}\check{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_k}}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{\sqrt{p_{-k}}}\right)\right), \quad \text{(B.36)}$$

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t}\left(\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}}\mathbf{\Delta}_k\mathbf{\Delta}_k^\top\right)\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top \mathbf{E}_{\chi,k}\check{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right)\right), \quad \text{(B.37)}$$

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t}\mathbf{\Delta}_k\mathbf{\Delta}_k^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}} - \mathbf{X}_{k,t})^\top \mathbf{E}_{\chi,k}\check{\mathbf{J}}_k^{[2]}\left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right\| = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\left(\frac{M_n}{\tau_{n,p}^{(k)}}\right)^{1+2\epsilon}\right). \quad \text{(B.38)}$$

*Proof.* For the proof of the first claim in (B.34), let us write

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| \le \frac{1}{np}\left\|\sum_{t\in[n]} \mathbf{Z}_{k,t}\left(\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}}\mathbf{\Delta}_k\mathbf{\Delta}_k^\top\right)\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\|$$

$$+\frac{1}{np}\left\|\frac{1}{p_{-k}}\sum_{t\in[n]}\mathbf{Z}_{k,t}\mathbf{\Delta}_k\mathbf{\Delta}_k^\top\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| =: U_1+U_2.$$

By Lemma B.17, we have

$$U_1^2\le\underbrace{\frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\left(\sum_{t\in[n]}Z_{k,i\ell,t}\mathsf{E}(X_{k,jm,t}^{\mathsf{t}})\right)^2}_{V_1}\left\|\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top-\frac{1}{p_{-k}}\right\|^2.$$

As shown in the proof of Lemma B.13 (i),

$$\mathsf{E}(V_1)\le\frac{c_\epsilon M_n^{2-2\epsilon}\omega^{2+2\epsilon}}{n}$$

which, together with Lemma B.18, leads to

$$U_1=O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}}\vee\frac{1}{p}\right)\right).$$

Also, we follow the steps in bounding $U_2$ appearing in the proof of Lemma B.13 (i), to show

$$U_2=O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right),$$

which completes the proof. The second claim therein as well as those in (B.35) follow analogously.

For the claim in (B.36), let us write

$$\begin{aligned}&\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}})\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top\mathbf{Z}_{k,t}^\top\mathbf{E}_{\chi,k}\right\|\\&\le\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}})\left(\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top-\frac{1}{p_{-k}}\mathbf{\Delta}_k\mathbf{\Delta}_k\right)\mathbf{Z}_{k,t}^\top\mathbf{E}_{\chi,k}\right\|\\&\quad+\frac{\sqrt{p_k}}{npp_{-k}}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}})\mathbf{\Delta}_k\mathbf{\Delta}_k\mathbf{Z}_{k,t}^\top\mathbf{E}_{\chi,k}\right\| =: U_3+U_4.\end{aligned}$$

By Lemma B.17, we have

$$U_3^2\le\underbrace{\frac{1}{n^2p_kp_{-k}^2}\sum_{q\in[r_k]}\sum_{\ell,m\in[p_{-k}]}\left(\sum_{j\in[p_k]}\sum_{t\in[n]}e_{\chi,jq}^{(k)}\mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}})Z_{k,jm,t}\right)^2}_{V_3}\left\|\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top-\frac{1}{p_{-k}}\mathbf{\Delta}_k\mathbf{\Delta}_k\right\|^2$$

where, by Lemmas B.1 (i), B.2, B.3 and B.24 (i) and Assumptions 3 (iii) and 4 (i),

$$\begin{aligned}\mathsf{E}(V_3) &= \frac{1}{n^2 p_k p_{-k}^2} \sum_{q\in[r_k]} \sum_{\ell,m\in[p_{-k}]} \sum_{j\in[p_k]} \sum_{t,u\in[n]} (e_{\chi,jq}^{(k)})^2 \mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}) \mathsf{E}(X_{k,i\ell,u}^{\mathsf{t}}) \mathsf{Cov}(Z_{k,jm,t}, Z_{k,jm,u}) \\ &\lesssim \frac{1}{n^2 p_k} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^2 (|\mathcal{F}_u|_2+\omega)^2 \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \le \frac{c_\epsilon M_n^{2-2\epsilon}\omega^{2+2\epsilon}}{np_k}.\end{aligned}$$

Then from Lemma B.18,

$$U_3 = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_k}}\left(\sum_{k'\in[K]\setminus\{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right)\right).$$

Similarly, additionally evoking Lemma B.10 (iv),

$$\begin{aligned}\mathsf{E}(U_4^2) &\le \frac{1}{n^2 p_k p_{-k}^4} \sum_{q\in[r_k]} \mathsf{E}\left[\left(\sum_{j\in[p_k]} \sum_{\ell,m\in[p_{-k}]} \sum_{q'\in[r_{-k}]} \sum_{t\in[n]} e_{\chi,jq}^{(k)} \delta_{k,\ell q'} \delta_{k,mq'} \mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}) Z_{k,jm,t}\right)^2\right] \\ &= \frac{1}{n^2 p_k p_{-k}^4} \sum_{q\in[r_k]} \sum_{j\in[p_k]} \sum_{\ell,\ell',m\in[p_{-k}]} \sum_{q',q''\in[r_{-k}]} \sum_{t,u\in[n]} (e_{\chi,jq}^{(k)})^2 \delta_{k,\ell q'} \delta_{k,mq'} \delta_{k,\ell' q''} \delta_{k,mq''} \\ &\quad \times \mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}) \mathsf{E}(X_{k,i\ell',u}^{\mathsf{t}}) \mathsf{Cov}(Z_{k,jm,t}, Z_{k,jm,u}) \\ &\lesssim \frac{1}{n^2 p} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^2 (|\mathcal{F}_u|_2+\omega)^2 \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \lesssim \frac{c_\epsilon M_n^{2-2\epsilon}\omega^{2+2\epsilon}}{np},\end{aligned}$$

which leads to

$$U_4 = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np}}\right).$$

Collecting the bounds on $U_3$ and $U_4$ and evoking Lemma A.1 and (B.26) completes the proof. For the claim in (B.37), let us write

$$U_5 := \frac{\sqrt{p_k}}{np} \left\| \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t} \left(\breve{\mathbf{D}}_k^{[1]} (\breve{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \mathbf{\Delta}_k \mathbf{\Delta}_k^\top\right) \mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top \mathbf{E}_{\chi,k} \right\|.$$

By Lemma B.17, we have

$$U_5^2 \le \underbrace{\frac{1}{n^2 p_k p_{-k}^2} \sum_{q\in[r_k]} \sum_{\ell,m\in[p_{-k}]} \left(\sum_{j\in[p_k]} \sum_{t\in[n]} e_{\chi,jq}^{(k)} Z_{k,i\ell,t} \mathsf{E}(X_{k,jm,t}^{\mathsf{t}})\right)^2}_{V_5} \left\| \breve{\mathbf{D}}_k^{[1]} (\breve{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \mathbf{\Delta}_k \mathbf{\Delta}_k^\top \right\|^2$$

where, by Lemmas B.1 (i), B.2, B.3 and B.24 (i) and Assumptions 3 (iii) and 4 (i),

$$
\begin{aligned}
\mathsf{E}(V_5) &= \frac{1}{n^2 p_k p_{-k}^2} \sum_{q\in[r_k]} \sum_{\ell,m\in[p_{-k}]} \sum_{j,j'\in[p_k]} \sum_{t,u\in[n]} e^{(k)}_{\chi,jq} e^{(k)}_{\chi,j'q} \mathsf{E}(X^{\mathsf{t}}_{k,jm,t}) \mathsf{E}(X^{\mathsf{t}}_{k,j'm,u}) \mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,i\ell,u}) \\
&\lesssim \frac{1}{n^2} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^2 (|\mathcal{F}_u|_2+\omega)^2 \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \\
&\le \frac{c_\epsilon M_n^{2-2\epsilon}\omega^{2+2\epsilon}}{n} \lesssim \left(\frac{M_n^{1-\epsilon}}{\sqrt{n}}\right)^2,
\end{aligned}
$$

which leads to

$$
U_5 = O_P\left[\frac{M_n^{1-\epsilon}}{\sqrt{n}} \left(\sum_{k'\in[K]\setminus\{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right)\right]
$$

by Markov's inequality and Lemma B.18. Combined with Lemma A.1, the proof is complete. Finally, for (B.38), we write

$$
U_6 := \frac{\sqrt{p_k}}{npp_{-k}} \left\| \sum_{t\in[n]} \mathbf{Z}_{k,i\cdot,t} \mathbf{\Delta}_k \mathbf{\Delta}_k^\top \mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t})^\top \mathbf{E}_{\chi,k} \right\|.
$$

By Lemmas B.1 (iii) and B.10 (iv) and those arguments adopted in bounding $U_5$,

$$
\begin{aligned}
\mathsf{E}(U_6^2) &\le \frac{1}{n^2 p_k p_{-k}^4} \sum_{q\in[r_k]} \mathsf{E}\left[\left(\sum_{j\in[p_k]} \sum_{\ell,m\in[p_{-k}]} \sum_{q'\in[r_{-k}]} \sum_{t\in[n]} e^{(k)}_{\chi,jq} \delta_{k,\ell q'} \delta_{k,mq'} \mathsf{E}(X^{\mathsf{t}}_{k,jm,t} - X_{k,jm,t}) Z_{k,i\ell,t}\right)^2\right] \\
&= \frac{1}{n^2 p_k p_{-k}^4} \sum_{q\in[r_k]} \sum_{j,j'\in[p_k]} \sum_{\ell,m,m'\in[p_{-k}]} \sum_{q',q''\in[r_{-k}]} \sum_{t,u\in[n]} e^{(k)}_{\chi,jq} e^{(k)}_{\chi,j'q} \delta_{k,\ell q'} \delta_{k,mq'} \delta_{k,\ell q''} \delta_{k,m'q''} \\
&\quad \times \mathsf{E}(X^{\mathsf{t}}_{k,jm,t} - X_{k,jm,t}) \mathsf{E}(X^{\mathsf{t}}_{k,j'm',u} - X_{k,j'm',u}) \mathsf{Cov}(Z_{k,i\ell,t}, Z_{k,i\ell,u}) \\
&\lesssim \frac{1}{n^2 p_{-k} (\tau^{(k)}_{n,p})^{2+4\epsilon}} \sum_{t,u\in[n]} (|\mathcal{F}_t|_2+\omega)^{3+2\epsilon} (|\mathcal{F}_u|_2+\omega)^{3+2\epsilon} \exp\left(-\frac{c_0\epsilon|t-u|}{1+\epsilon}\right) \\
&\lesssim \frac{c_\epsilon \omega^{2+2\epsilon} M_n^{4+2\epsilon}}{np_{-k} (\tau^{(k)}_{n,p})^{2+4\epsilon}} \lesssim \left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \left(\frac{M_n}{\tau^{(k)}_{n,p}}\right)^{1+2\epsilon}\right)^2,
\end{aligned}
$$

which leads to

$$
U_6 = O_P\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \left(\frac{M_n}{\tau^{(k)}_{n,p}}\right)^{1+2\epsilon}\right).
$$

Evoking Lemma A.1 and (B.26) completes the proof of (B.38). □

**Lemma B.21.** *For each $k \in [K]$, we have*

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}\right) \check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t}\right)^\top\right\| = O_P\left(\frac{M_n^\epsilon \sqrt{\log(np_{-k})}}{(\tau^{(k)}_{n,p})^\epsilon} \cdot \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right), \quad \text{(B.39a)}$$

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}_{k,t}\right) \check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}\right)^\top\right\| = O_P\left(\frac{M_n^\epsilon \sqrt{\log(np_{-k})}}{(\tau^{(k)}_{n,p})^\epsilon} \cdot \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right). \quad \text{(B.39b)}$$

*Also, for any $i \in [p_k]$,*

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,i\cdot,t} - \mathbf{X}_{k,i\cdot,t}\right) \check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t}\right)^\top\right\| = O_P\left(\frac{M_n^\epsilon \sqrt{\log(np_{-k})}}{(\tau^{(k)}_{n,p})^\epsilon} \cdot \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right), \quad \text{(B.40a)}$$

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}_{k,i\cdot,t}\right) \check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}\right)^\top\right\| = O_P\left(\frac{M_n^\epsilon \sqrt{\log(np_{-k})}}{(\tau^{(k)}_{n,p})^\epsilon} \cdot \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right), \quad \text{(B.40b)}$$

*and*

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,i\cdot,t} - \mathbf{X}_{k,i\cdot,t}\right) \check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t}\right)^\top \mathbf{E}_{\chi,k} \check{\mathbf{J}}^{[2]}_k \left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right\|$$
$$= O_P\left(\frac{M_n^\epsilon \sqrt{\log(np_{-k})}}{(\tau^{(k)}_{n,p})^\epsilon} \cdot \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right), \quad \text{(B.41a)}$$

$$\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]} \mathsf{E}\left(\mathbf{X}_{k,i\cdot,t}\right) \check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}\left(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}\right)^\top \mathbf{E}_{\chi,k} \check{\mathbf{J}}^{[2]}_k \left(\frac{1}{p_k}\mathbf{M}_{\chi,k}\right)^{-1}\right\|$$
$$= O_P\left(\frac{M_n^\epsilon \sqrt{\log(np_{-k})}}{(\tau^{(k)}_{n,p})^\epsilon} \cdot \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right). \quad \text{(B.41b)}$$

*Proof.* For the proof of (B.39), let us write

$$\frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}) \check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top \mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t})^\top\right\|$$
$$\le \frac{1}{np}\left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t} - \mathbf{X}_{k,t}) \left(\check{\mathbf{D}}^{[1]}_k (\check{\mathbf{D}}^{[1]}_k)^\top - \frac{1}{p_{-k}}\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top\right) \mathsf{E}(\mathbf{X}^{\mathsf{t}}_{k,t})^\top\right\|$$

$$+\frac{1}{np_kp_{-k}^2}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}}-\mathbf{X}_{k,t})\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\right\| =: U_1+U_2.$$

By Lemma B.17, we have

$$U_1^2 \le \frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\left(\sum_{t\in[n]}\mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}-X_{k,i\ell,t})\mathsf{E}(X_{k,jm,t}^{\mathsf{t}})\right)^2\left\|\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top-\frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right\|^2,$$

where the proof of Lemma B.14 and Lemma B.18 show that

$$U_1 = O_P\left(\frac{M_n}{\tau_{n,p}^{(k)}}\psi_{n,p}^{(k)}\left(\sum_{k'\in[K]\setminus\{k\}}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}}\vee\frac{1}{p}\right)\right).$$

Also, proceeding similarly as in the arguments in the proof of Lemma B.14 for bounding $U_2$ therein, we obtain

$$U_2 = O\left(\frac{M_n}{\tau_{n,p}^{(k)}}\psi_{n,p}^{(k)}\right) = O_P\left(\frac{M_n^\epsilon\log(np_{-k})}{(\tau_{n,p}^{(k)})^\epsilon}\cdot\frac{1}{\sqrt{np_{-k}}}\right),$$

which completes the proof of the first claim. The second claim in (B.39) as well as those in (B.40) are proved following the analogous steps.

For the the proof of (B.41), let us write

$$\begin{aligned}
&\frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}}-\mathbf{X}_{k,i\cdot,t})\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\mathbf{E}_{\chi,k}\right\| \\
&\le \frac{\sqrt{p_k}}{np}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}}-\mathbf{X}_{k,i\cdot,t})\left(\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top-\frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right)\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\mathbf{E}_{\chi,k}\right\| \\
&\quad+\frac{\sqrt{p_k}}{npp_{-k}}\left\|\sum_{t\in[n]}\mathsf{E}(\mathbf{X}_{k,i\cdot,t}^{\mathsf{t}}-\mathbf{X}_{k,i\cdot,t})\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\mathsf{E}(\mathbf{X}_{k,t}^{\mathsf{t}})^\top\mathbf{E}_{\chi,k}\right\| =: U_3+U_4.
\end{aligned}$$

By Lemma B.17, we have

$$\begin{aligned}
U_3^2 \le \underbrace{\frac{1}{n^2p_kp_{-k}^2}\sum_{q\in[r_k]}\sum_{\ell,m\in[p_{-k}]}\left(\sum_{j\in[p_k]}\sum_{t\in[n]}e_{\chi,jq}^{(k)}\mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}}-X_{k,i\ell,t})\mathsf{E}(X_{k,jm,t}^{\mathsf{t}})\right)^2}_{V_3} \\
\times\left\|\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top-\frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right\|^2
\end{aligned}$$

where, by Lemmas B.1 (i), (iii) and B.24 (i) and Assumption 3 (ii) and (iii),

$$V_3 = \frac{1}{n^2 p_k p_{-k}^2} \sum_{q\in[r_k]} \sum_{\ell,m\in[p_{-k}]} \sum_{j,j'\in[p_k]} \sum_{t,u\in[n]} e_{\chi,jq}^{(k)} e_{\chi,j'q}^{(k)}$$
$$\times \mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}} - X_{k,i\ell,t})\mathsf{E}(X_{k,i\ell,u}^{\mathsf{t}} - X_{k,i\ell,u})\mathsf{E}(X_{k,jm,t}^{\mathsf{t}})\mathsf{E}(X_{k,j'm,u}^{\mathsf{t}})$$
$$\lesssim \left(\frac{1}{n(\tau_{n,p}^{(k)})^{1+2\epsilon}} \sum_{t\in[n]} (|\mathcal{F}_t|_2 + \omega)^{3+2\epsilon}\right)^2 \lesssim \left(\frac{M_n \omega^{2+2\epsilon}}{(\tau_{n,p}^{(k)})^{1+2\epsilon}}\right)^2 = \left(\frac{M_n}{\tau_{n,p}^{(k)}} \psi_{n,p}^{(k)}\right)^2$$

which, together with Lemma B.18, leads to

$$U_3 = O_P\left(\frac{M_n}{\tau_{n,p}^{(k)}} \psi_{n,p}^{(k)} \left(\sum_{k'\in[K]\setminus\{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right)\right).$$

Similarly, additionally evoking Lemma B.10 (iv),

$$U_4^2 \le \frac{1}{n^2 p_k p_{-k}^4} \sum_{q\in[r_k]} \left(\sum_{j\in[p_k]} \sum_{\ell,m\in[p_{-k}]} \sum_{q'\in[r_{-k}]} \sum_{t\in[n]} e_{\chi,jq}^{(k)} \delta_{k,\ell q'} \delta_{k,mq'} \mathsf{E}(X_{k,i\ell,t}^{\mathsf{t}} - X_{k,i\ell,t})\mathsf{E}(X_{k,jm,t}^{\mathsf{t}})\right)^2$$
$$\lesssim \left(\frac{1}{n(\tau_{n,p}^{(k)})^{1+2\epsilon}} \sum_{t\in[n]} (|\mathcal{F}_t|_2 + \omega)^{3+2\epsilon}\right)^2 \lesssim \left(\frac{M_n}{\tau_{n,p}^{(k)}} \psi_{n,p}^{(k)}\right)^2 \lesssim \left(\frac{M_n^{\epsilon}\sqrt{\log(np_{-k}}}{(\tau_{n,p}^{(k)})^{\epsilon}} \cdot \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right)^2,$$

leading to

$$U_4 = O\left(\frac{M_n}{\tau_{n,p}^{(k)}} \psi_{n,p}^{(k)}\right).$$

Collecting the bounds on $U_3$ and $U_4$ and evoking Lemma A.1 and (B.26) completes the proof. The remaining claim is proved following the analogous steps. □

**Lemma B.22.** *For each $k \in [K]$, we have*

$$\left\|\frac{1}{np} \sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}) \left(\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top\right) \mathsf{E}(\mathbf{X}_{k,t})^\top\right\| = O_P\left(\sum_{k'\in[K]\setminus\{k\}} \frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}} \vee \frac{1}{p}\right).$$

*Proof.* By Lemma B.17, we have

$$\frac{1}{n^2 p^2} \left\|\sum_{t\in[n]} \mathsf{E}(\mathbf{X}_{k,t}) \left(\check{\mathbf{D}}_k^{[1]}(\check{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}} \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^\top\right) \mathsf{E}(\mathbf{X}_{k,t})^\top\right\|_F^2$$

$$\leq \underbrace{\frac{1}{n^2p^2}\sum_{i,j\in[p_k]}\sum_{\ell,m\in[p_{-k}]}\left(\sum_{t\in[n]}\mathsf{E}(X^{\mathsf{t}}_{k,i\ell,t})\mathsf{E}(X^{\mathsf{t}}_{k,jm,t})\right)^2}_{U_1}\left\|\breve{\mathbf{D}}_k^{[1]}(\breve{\mathbf{D}}_k^{[1]})^\top - \frac{1}{p_{-k}}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^\top\right\|_F^2.$$

By Lemma B.1 (i) and Assumption 3 (ii), we have

$$U_1 \lesssim \left(\frac{1}{n}\sum_{t\in[n]}(|\mathcal{F}_t|_2+\omega)^2\right)^2 \lesssim \omega^4$$

which, in combination with Lemma B.18, leads to the claim. □

**Lemma B.23.** *Let the conditions in Theorem 3 hold. For $\breve{\mathbf{M}}_k^{[2]}(\tau) \in \mathbb{R}^{r_k\times r_k}$ denoting the diagonal matrix containing the eigenvalues $\breve{\mu}_j^{(k),[2]}(\tau)$, $j\in[r_k]$, of $\breve{\boldsymbol{\Gamma}}^{(k),[2]}(\tau)$ on its diagonal, we have*

$$\left\|\left(p_k^{-1}\breve{\mathbf{M}}_k^{[2]}(\tau)\right)^{-1} - \left(p_k^{-1}\mathbf{M}_{\chi,k}\right)^{-1}\right\| = O_P\left(\sum_{k'\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}}\vee\frac{1}{p}\right).$$

*Proof.* The proof proceeds analogously as in Lemma B.16 with (B.24) in place of Proposition B.2, and thus is omitted. □

**Lemma B.24.** *Let the conditions in Theorem 3 hold.*

*(i)* $|\mathbf{E}_{\chi,k}|_\infty \leq Cp_k^{-1/2}$ *with some constant* $C\in(0,\infty)$.

*(ii) There exists a diagonal matrix* $\breve{\mathbf{J}}_k^{[2]}\in\mathbb{R}^{r_k\times r_k}$ *with* $\pm1$ *on its diagonal, such that*

$$\left\|\breve{\mathbf{E}}_k^{[2]}(\tau) - \mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\right\| = O_P\left(\sum_{k'\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}}\vee\frac{1}{p}\right).$$

*(iii) With* $\breve{\mathbf{J}}_k^{[2]}\in\mathbb{R}^{r_k\times r_k}$ *from (ii), it holds that*

$$\left\|\breve{\mathbf{E}}_k^{[2]}(\tau)\left(p_k^{-1}\breve{\mathbf{M}}_k^{[2]}(\tau)\right)^{-1} - \mathbf{E}_{\chi,k}\breve{\mathbf{J}}_k^{[2]}\left(p_k^{-1}\mathbf{M}_{\chi,k}\right)^{-1}\right\| = O_P\left(\sum_{k'\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k'}}}\vee\frac{1}{p}\right).$$

*(iv) There exists a diagonal matrix* $\bar{\mathbf{J}}_k\in\mathbb{R}^{r_k\times r_k}$ *with* $\pm1$ *on its diagonal, such that we have* $p_k^{-1/2}\mathbf{E}_{\chi,k}^\top\boldsymbol{\Lambda}_k = \bar{\mathbf{J}}_k + o(1)$.

*Proof.* Firstly, note that

$$|\mathbf{\Gamma}_\chi^{(k)}|_\infty \le r_k \bar{\lambda}^2 \omega^2 + o(1)$$

under Assumptions 1, 2 and 3 (ii). Then, due to (B.24), the conditions (C1)–(C3) are met with $\mathbf{S} = \widetilde{\mathbf{S}} = \mathbf{\Gamma}_\chi^{(k)}$ and $\widehat{\mathbf{S}} = \check{\mathbf{\Gamma}}^{(k),[2]}$, which leads to (i) and (ii) from Lemma B.4 (i) and (iii). As for (iii), since

$$\begin{aligned}
&\left\| \check{\mathbf{E}}_k^{[2]}(\tau) \left( p_k^{-1} \widecheck{\mathbf{M}}_k^{[2]}(\tau) \right)^{-1} - \mathbf{E}_{\chi,k} \check{\mathbf{J}}_k^{[2]} \left( p_k^{-1} \mathbf{M}_{\chi,k} \right)^{-1} \right\| \\
\le & \left\| \check{\mathbf{E}}_k^{[2]}(\tau) - \mathbf{E}_{\chi,k} \check{\mathbf{J}}_k^{[2]} \right\| \left\| \left( p_k^{-1} \widecheck{\mathbf{M}}_k^{[2]}(\tau) \right)^{-1} \right\| + \|\mathbf{E}_{\chi,k}\| \left\| \left( p_k^{-1} \widecheck{\mathbf{M}}_k^{[2]}(\tau) \right)^{-1} - \left( p_k^{-1} \mathbf{M}_{\chi,k} \right)^{-1} \right\|,
\end{aligned}$$

the claim follows from Lemma B.23 and (B.25). Finally, from repeated application of Assumptions 1 and 2,

$$\begin{aligned}
\frac{1}{\sqrt{p_k}} \mathbf{\Lambda}_k^\top \mathbf{E}_{\chi,k} \left( \frac{1}{p_k} \mathbf{M}_{\chi,k} \right) &= \frac{1}{p_k^{3/2}} \mathbf{\Lambda}_k^\top \mathbf{\Gamma}_\chi^{(k)} \mathbf{E}_{\chi,k} \\
&= \frac{1}{p_k} \mathbf{\Lambda}_k^\top \mathbf{\Lambda}_k \cdot \frac{1}{n} \sum_{t \in [n]} \mathsf{mat}_k(\mathcal{F}_t) \mathsf{mat}_k(\mathcal{F}_t)^\top \left( \frac{1}{\sqrt{p_k}} \mathbf{\Lambda}_k^\top \mathbf{E}_{\chi,k} \right) \\
&= \mathbf{\Gamma}_f^{(k)} \left( \frac{1}{\sqrt{p_k}} \mathbf{\Lambda}_k^\top \mathbf{E}_{\chi,k} \right) + o(1).
\end{aligned}$$

Evoking Lemma A.1 , we have $p_k^{-1/2} \mathbf{\Lambda}_k^\top \mathbf{E}_{\chi,k} = \bar{\mathbf{J}}_k + o(1)$. □

## B.5 Proof of Theorem 4

For notational convenience, we omit the dependence on $\tau$ and $\iota$, where $\iota = 2$ under Assumption 4 and $\iota = 1$ under Assumption 5 throughout the proof.

### B.5.1 Proof of (14)

Let define $\mathbf{\Lambda} := \otimes_{k=K}^1 \mathbf{\Lambda}_k$, $\check{\mathbf{\Lambda}} := \otimes_{k=K}^1 \check{\mathbf{\Lambda}}_k$ and $\check{\mathbf{H}} := \otimes_{k=K}^1 \check{\mathbf{H}}_k$. Then, it follows that $\check{\mathbf{H}}^\top \check{\mathbf{H}} = \mathbf{I}_r + o_P(1)$ such that $\check{\mathbf{H}}$ is asymptotically invertible. Combining this with Assumption 1, Theorems 2 (under Assumption 5) and 3 (under Assumption 4), we have

$$\frac{1}{\sqrt{p}} \left\| \check{\mathbf{\Lambda}} - \mathbf{\Lambda} \check{\mathbf{H}} \right\| = \begin{cases} O_P\left( \sum_{k \in [K]} \frac{M_n^{1-\epsilon}}{\sqrt{n p_{-k}}} \vee \frac{1}{p} \right) & \text{under Assumption 4,} \\ O_P\left( \sum_{k \in [K]} \frac{M_n^{1-\epsilon}}{\sqrt{n p_{-k}}} \vee \frac{1}{p} \vee \bar{\psi}_{n,p} \right) & \text{under Assumption 5,} \end{cases} \tag{B.42}$$

by the arguments adopted in Lemma B.10 (i). Further, the same error bounded is inherited to $\|\breve{\mathbf{\Lambda}}\breve{\mathbf{H}}^{-1} - \mathbf{\Lambda}\|$, as

$$\frac{1}{\sqrt{p}}\left\|\breve{\mathbf{\Lambda}}\breve{\mathbf{H}}^{-1} - \mathbf{\Lambda}\right\| = \begin{cases} O_P\left(\sum_{k\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p}\right) & \text{under Assumption 4,} \\ O_P\left(\sum_{k\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p} \vee \bar{\psi}_{n,p}\right) & \text{under Assumption 5,} \end{cases} \tag{B.43}$$

Next, noting that $\mathsf{vec}(\widehat{\mathcal{F}}_t) = p^{-1}\breve{\mathbf{\Lambda}}^\top \mathsf{vec}(\mathcal{X}_t^{\mathsf{t}})$, we write

$$\mathsf{vec}(\widehat{\mathcal{F}}_t) - \breve{\mathbf{H}}^{-1}\mathsf{vec}(\mathcal{F}_t) = \frac{1}{p}\breve{\mathbf{\Lambda}}^\top\left(\mathbf{\Lambda} - \breve{\mathbf{\Lambda}}\widehat{\mathbf{H}}^{-1}\right)\mathsf{vec}\left(\mathcal{F}_t\right) + \frac{1}{p}\breve{\mathbf{\Lambda}}^\top\mathsf{vec}(\boldsymbol{\xi}_t) =: U_1 + U_2. \tag{B.44}$$

Then by (B.43),

$$\begin{aligned}|U_1|_2 &\le \frac{1}{p}\|\breve{\mathbf{\Lambda}}\|\left\|\mathbf{\Lambda} - \breve{\mathbf{\Lambda}}\widehat{\mathbf{H}}^{-1}\right\||\mathcal{F}_t|_2 \\ &= \begin{cases} O_P\left(|\mathcal{F}_t|_2\cdot\sum_{k\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p}\right) & \text{under Assumption 4,} \\ O_P\left(|\mathcal{F}_t|_2\cdot\sum_{k\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p} \vee \bar{\psi}_{n,p}\right) & \text{under Assumption 5.} \end{cases}\end{aligned}$$

Also, writing

$$U_2 = \frac{1}{p}\left(\breve{\mathbf{\Lambda}} - \mathbf{\Lambda}\breve{\mathbf{H}}\right)^\top\mathsf{vec}(\boldsymbol{\xi}_t) + \frac{1}{p}\breve{\mathbf{H}}^\top\mathbf{\Lambda}^\top\mathsf{vec}(\boldsymbol{\xi}_t) =: V_{2,1} + V_{2,2},$$

we have from (B.42),

$$\begin{aligned}|V_{2,1}|_2 &\le \frac{1}{p}\left\|\breve{\mathbf{\Lambda}} - \mathbf{\Lambda}\breve{\mathbf{H}}\right\||\mathsf{vec}(\boldsymbol{\xi}_t)|_2 \\ &= \begin{cases} O_P\left(\omega\sum_{k\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p}\right) & \text{under Assumption 4,} \\ O_P\left(\omega\sum_{k\in[K]}\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}} \vee \frac{1}{p} \vee \bar{\psi}_{n,p}\right) & \text{under Assumption 5,} \end{cases}\end{aligned}$$

from the observation that under Assumption 3 (i),

$$\mathsf{E}\left(|\mathsf{vec}(\boldsymbol{\xi}_t)|_2^2\right) \le p\omega^2. \tag{B.45}$$

With some notational abuse, denote by $\mathbf{\Lambda} = [\lambda_{\mathbf{ij}}, \mathbf{i} \in \prod_{k=1}^K [p_k], \mathbf{j} \in \prod_{k=1}^K [r_k]]$. Under Assumptions 4, we have

$$\begin{aligned}\frac{1}{p^2}\mathsf{E}\left(\left|\mathbf{\Lambda}^\top\mathsf{vec}(\boldsymbol{\xi}_t)\right|_2^2\right) &= \frac{1}{p^2}\sum_{\mathbf{j}\in\prod_{k=1}^K[r_k]}\mathsf{E}\left(\left|\sum_{\mathbf{i}\in\prod_{k=1}^K[p_k]}\lambda_{\mathbf{ij}}\xi_{\mathbf{i},t}\right|^2\right) \\ &= \frac{1}{p^2}\sum_{\mathbf{j}\in\prod_{k=1}^K[r_k]}\sum_{\mathbf{i}\in\prod_{k=1}^K[p_k]}\lambda_{\mathbf{ij}}^2\mathsf{E}(\xi_{\mathbf{i},t}^2) \lesssim \frac{r\omega^2}{p}\end{aligned}$$

by Assumption 1 (ii), which leads to $|V_{2,2}|_2 = O_P(\omega p^{-1/2})$. Under Assumptions 5,

$$\begin{aligned}\frac{1}{p^2}\mathsf{E}\left(\left|\boldsymbol{\Lambda}^\top \mathsf{vec}(\boldsymbol{\xi}_t)\right|_2^2\right) &= \frac{1}{p^2}\sum_{\mathbf{j}\in\prod_{k=1}^K[r_k]}\sum_{\mathbf{i},\mathbf{i}'\in\prod_{k=1}^K[p_k]}\lambda_{\mathbf{ij}}\lambda_{\mathbf{i}'\mathbf{j}}\mathsf{Cov}(\xi_{\mathbf{i},t},\xi_{\mathbf{i}',t})\\
&\lesssim \frac{r}{p^2}\sum_{\mathbf{i},\mathbf{i}'\in\prod_{k=1}^K[p_k]}\|\xi_{\mathbf{i},t}\|_\nu\|\xi_{\mathbf{i}',t}\|_\nu \exp\left(-\frac{c_0(\nu-2)|\mathbf{i}-\mathbf{i}'|_2}{K\nu}\right)\\
&\lesssim \frac{r\omega^2}{p}\prod_{k=1}^K\frac{1}{p_k}\sum_{i_k,i_k'\in[p_K]}\exp\left(-\frac{c_0\epsilon|i_k-i_k'|}{K(1+\epsilon)}\right)\lesssim\frac{r\omega^2}{p}\end{aligned}$$

with $\nu = 2+2\epsilon$, due to Lemma B.3, and thus we have $|V_{2,2}|_2 = O_P(\omega p^{-1/2})$.

### B.5.2 Proof of (15)

From the proof of (14),

$$\begin{aligned}&\frac{1}{n}\sum_{t\in[n]}\left|\mathsf{vec}(\widehat{\mathcal{F}}_t) - \breve{\mathbf{H}}^{-1}\mathsf{vec}(\mathcal{F}_t)\right|_2^2\\
\le&\frac{2}{n}\sum_{t\in[n]}\left|\frac{1}{p}\breve{\boldsymbol{\Lambda}}^\top\left(\boldsymbol{\Lambda}-\breve{\boldsymbol{\Lambda}}\widehat{\mathbf{H}}^{-1}\right)\mathsf{vec}\left(\mathcal{F}_t\right)\right|_2^2 + \frac{2}{n}\sum_{t\in[n]}\left|\frac{1}{p}\breve{\boldsymbol{\Lambda}}^\top\mathsf{vec}(\boldsymbol{\xi}_t)\right|_2^2 =: U_1+U_2.\end{aligned}$$

Then by (B.43) and Assumption 3 (ii),

$$\begin{aligned}|U_1| &\lesssim \frac{1}{p^2}\|\breve{\boldsymbol{\Lambda}}\|^2\left\|\boldsymbol{\Lambda}-\breve{\boldsymbol{\Lambda}}\widehat{\mathbf{H}}^{-1}\right\|^2\cdot\frac{1}{n}\sum_{t\in[n]}|\mathcal{F}_t|_2^2\\
&=\begin{cases}O_P\left[\omega^2\sum_{k\in[K]}\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right)^2\vee\frac{1}{p^2}\right] & \text{under Assumption 4,}\\ O_P\left[\omega^2\cdot\sum_{k\in[K]}\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right)^2\vee\frac{1}{p^2}\vee\bar{\psi}_{n,p}^2\right] & \text{under Assumption 5.}\end{cases}\end{aligned}$$

As for $U_2$, by the arguments analogous to those adopted in the proof of (14),

$$\begin{aligned}U_2&\lesssim\frac{1}{n}\sum_{t\in[n]}|\mathsf{vec}(\boldsymbol{\xi}_t)|_2^2\cdot\left\|\frac{1}{p}\left(\breve{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}\breve{\mathbf{H}}\right)^\top\right\| + \frac{1}{n}\sum_{t\in[n]}\left\|\frac{1}{p}\boldsymbol{\Lambda}^\top\mathsf{vec}(\boldsymbol{\xi}_t)\right\|^2\\
&=\begin{cases}O_P\left[\omega^2\sum_{k\in[K]}\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right)^2\vee\frac{1}{p}\right] & \text{under Assumption 4,}\\ O_P\left[\omega^2\sum_{k\in[K]}\left(\frac{M_n^{1-\epsilon}}{\sqrt{np_{-k}}}\right)^2\vee\frac{1}{p}\vee\bar{\psi}_{n,p}^2\right) & \text{under Assumption 5,}\end{cases}\end{aligned}$$

which completes the proof.

### B.5.3 Proof of Theorem 4 (ii)

Due the condition that $M_n = M$ and $\max_{k\in[K]} p_k = o(n)$, we have $\sqrt{p} = o(\min_{k\in[K]} \sqrt{np_{-k}})$. From Lemma B.24 (ii), analogously as in Lemma B.10 (i), we have

$$\left\|\breve{\mathbf{E}} - \mathbf{E}_\chi \breve{\mathbf{J}}\right\| = O_P\left(\sum_{k\in[K]} \frac{1}{\sqrt{np_{-k}}} \vee \frac{1}{p}\right), \tag{B.46}$$

where $\mathbf{E}_\chi := \mathbf{E}_{\chi,K} \otimes \ldots \otimes \mathbf{E}_{\chi,1}$, $\breve{\mathbf{E}} := \breve{\mathbf{E}}_K \otimes \ldots \otimes \breve{\mathbf{E}}_1$, and $\breve{\mathbf{J}} := \breve{\mathbf{J}}_K \otimes \ldots \otimes \breve{\mathbf{J}}_1 \in \mathbb{R}^{r\times r}$, denotes a diagonal matrix with $\pm 1$ on its diagonal. Also, from Lemma B.24 (i), it follows that

$$|\mathbf{E}_\chi|_\infty \lesssim p^{-1/2}. \tag{B.47}$$

As in (B.44), we write

$$\begin{aligned}\sqrt{p}\left(\mathsf{vec}(\widehat{\mathcal{F}}_t) - \breve{\mathbf{H}}^{-1}\mathsf{vec}(\mathcal{F}_t)\right) &= \frac{1}{\sqrt{p}}\breve{\mathbf{\Lambda}}^\top\left(\mathbf{\Lambda} - \breve{\mathbf{\Lambda}}\breve{\mathbf{H}}^{-1}\right)\mathsf{vec}(\mathcal{F}_t) + \left(\breve{\mathbf{E}} - \mathbf{E}_\chi\breve{\mathbf{J}}\right)^\top \mathsf{vec}(\boldsymbol{\xi}_t) \\ &\quad + \breve{\mathbf{J}}\mathbf{E}_\chi^\top \mathsf{vec}(\boldsymbol{\xi}_t) =: U_1 + U_2 + \breve{\mathbf{J}}\mathbf{E}_\chi^\top\mathsf{vec}(\boldsymbol{\xi}_t).\end{aligned}$$

By the arguments analogous to those adopted in the proof of (14), we have $U_1 = o_P(1)$. Also, from (B.45) and (B.46),

$$U_2 = \left(\breve{\mathbf{E}} - \mathbf{E}_\chi\breve{\mathbf{J}}\right)^\top \mathsf{vec}(\boldsymbol{\xi}_t) = O_P\left[\omega\left(\sum_{k\in[K]}\frac{\sqrt{p}}{\sqrt{np_{-k}}} \vee \frac{1}{\sqrt{p}}\right)\right] = o_P(1).$$

The leading term, $\breve{\mathbf{J}}\mathbf{E}_\chi^\top\mathsf{vec}(\boldsymbol{\xi}_t)$, has the covariance matrix

$$\mathbf{\Upsilon}_t := \mathsf{Cov}\left(\breve{\mathbf{J}}\mathbf{E}_\chi^\top\mathsf{vec}(\boldsymbol{\xi}_t)\right) = \mathbf{E}_\chi^\top\mathsf{Cov}\left(\mathsf{vec}(\boldsymbol{\xi}_t)\right)\mathbf{E}_\chi, \tag{B.48}$$

such that under Assumption 3 (i),

$$\|\mathbf{\Upsilon}_t\| \le \omega^2.$$

With some abuse of notation, let us write $\mathbf{E}_\chi = [e_{\chi,\mathbf{ij}}, \mathbf{i} \in \prod_{k=1}^K [p_k], \mathbf{j} \in \prod_{k=1}^K [r_k]]$. Then,

$$\mathbf{E}_\chi^\top\mathsf{vec}(\boldsymbol{\xi}_t) = \sum_{\mathbf{i}\in\prod_{k=1}^K[p_k]}\left(e_{\chi,\mathbf{ij}}\xi_{\mathbf{i},t}, \mathbf{j} \in \prod_{k=1}^K [r_k]\right)^\top =: \sum_{\mathbf{i}\in\prod_{k=1}^K[p_k]} \mathbf{Y}_{\mathbf{i},t}.$$

For any $\mathbf{a} = (a_{\mathbf{j}},\, j \in \prod_{k=1}^K [r_k]])^\top$ with $|\mathbf{a}|_2 = 1$, due to Assumption 4,

$$\sum_{\mathbf{i} \in \prod_{k=1}^K [p_k]} \mathsf{Var}(\mathbf{a}^\top \mathbf{Y}_{\mathbf{i},t}) = \mathsf{Var}\left(\mathbf{a}^\top \mathbf{E}_\chi^\top \mathsf{vec}(\boldsymbol{\xi}_t)\right) = \mathsf{Var}\left(\mathbf{a}^\top \mathbf{E}_\chi^\top \mathsf{vec}(\boldsymbol{\xi}_t)\right) \ge c_\xi \mathbf{a}^\top \mathbf{E}_\chi^\top \mathbf{E}_\chi \mathbf{a} = c_\xi,$$

where $c_\xi \in (0, \infty)$ is a constant satisfying $\min_{\mathbf{i} \in \prod_{k=1}^K [p_k]} \mathsf{Var}(\xi_{\mathbf{i},t}) \ge c_\xi$. Also thanks to (B.47),

$$\sum_{\mathbf{i} \in \prod_{k=1}^K [p_k]} \|\mathbf{a}^\top \mathbf{Y}_{\mathbf{i},t}\|_{2+2\epsilon}^{2+2\epsilon} = \sum_{\mathbf{i} \in \prod_{k=1}^K [p_k]} \left| \left( \sum_{\mathbf{j} \in \prod_{k=1}^K [r_k]} a_{\mathbf{j}} e_{\chi,\mathbf{ij}} \right) \right|^{2+2\epsilon} \|\xi_{\mathbf{i},t}\|_{2+2\epsilon}^{2+2\epsilon} \lesssim \omega^{2+2\epsilon} p^{-\epsilon}$$

under Assumption 3 (i). Altogether, we have

$$\frac{\sum_{\mathbf{i} \in \prod_{k=1}^K [p_k]} \|\mathbf{a}^\top \mathbf{Y}_{\mathbf{i},t}\|_{2+2\epsilon}^{2+2\epsilon}}{\left[\sum_{\mathbf{i} \in \prod_{k=1}^K [p_k]} \mathsf{Var}(\mathbf{a}^\top \mathbf{Y}_{\mathbf{i},t})\right]^{1+\epsilon}} \lesssim \frac{\omega^{2+2\epsilon} p^{-\epsilon}}{c_\xi^{1+\epsilon}} \to 0$$

as $\min(p_1, \ldots, p_K) \to \infty$. Hence, for any $\mathbf{a} \in \mathbb{R}^r$ with $|\mathbf{a}|_2 = 1$, we have

$$\sum_{\mathbf{i} \in \prod_{k=1}^K [p_k]} \mathbf{a}^\top \mathbf{Y}_{\mathbf{i},t} \to \mathcal{N}_r(\mathbf{0}, \mathbf{a}^\top \boldsymbol{\Upsilon}_t \mathbf{a})$$

by the Lyapunov condition (cf. Theorems 27.3 and 29.4 of Billingsley, 1995) which, in combination with Cramér-Wold theorem, leads to

$$\mathbf{E}_\chi^\top \mathsf{vec}(\boldsymbol{\xi}_t) \to_d \mathcal{N}_r(\mathbf{0}, \boldsymbol{\Upsilon}_t)$$

as $\min(p_1, \ldots, p_K) \to \infty$. This, combined with the bounds on $U_1$ and $U_2$ and that $\check{\mathbf{J}}$ is a diagonal matrix of signs, completes the proof.

### B.6 Proof of Proposition 1

Below, we omit the dependence on the truncation parameter $\tau$ where there is no confusion. The proof takes analogous steps as the proof of Theorem 3.9 of Barigozzi et al. (2022). Throughout, we denote the $j$-th largest eigenvalue of a square matrix $\mathbf{A}$ by $\lambda_j(\mathbf{A})$.
Note that for $\widehat{r}_k^{(m-1)} > r_k$, it holds that for

$$\widehat{\mathbf{D}}_k^{(m)} = \widehat{\mathbf{E}}_K^{(m)} \otimes \cdots \otimes \widehat{\mathbf{E}}_{k+1}^{(m)} \otimes \widehat{\mathbf{E}}_{k-1}^{(m)} \otimes \cdots \otimes \widehat{\mathbf{E}}_1^{(m)},$$

we have $\widehat{\mathbf{D}}_k^{(m)} (\widehat{\mathbf{D}}_k^{(m)})^\top - \widehat{\mathbf{D}}_k (\widehat{\mathbf{D}}_k)^\top$ non-negative definite since the difference is composed of the outer products of non-negative definite matrices, which in turn follows from that $\widehat{\mathbf{E}}_k^{(m)} =$

$[\widehat{\mathbf{E}}_k, \widetilde{\mathbf{E}}_k^{(m)}]$ with $\widehat{\mathbf{E}}_k^\top \widetilde{\mathbf{E}}_k^{(m)} = \mathbf{O}$. Then, noting that

$$\breve{\mathbf{\Gamma}}^{(k),(m)} = \breve{\mathbf{\Gamma}}^{(k)} + \frac{1}{np_{-k}} \sum_{t=1}^{n} \mathbf{X}_{k,t}^{\mathsf{t}} \left( \widehat{\mathbf{D}}_k^{(m)} (\widehat{\mathbf{D}}_k^{(m)})^\top - \widehat{\mathbf{D}}_k (\widehat{\mathbf{D}}_k)^\top \right) (\mathbf{X}_{k,t}^{\mathsf{t}})^\top,$$

we have for all $j \in [r_k]$,

$$\frac{1}{p_k} \lambda_j \left( \breve{\mathbf{\Gamma}}^{(k),(m)} \right) \geq \frac{1}{p_k} \lambda_j \left( \breve{\mathbf{\Gamma}}^{(k)} \right) \geq \alpha_j^{(k)} + O_P(\rho_{n,p}), \tag{B.49}$$

where the first inequality follows from Weyl's inequality, and the second from Assumption 2, Proposition B.2 and the condition on $\rho_{n,p}$. Also, since

$$\widehat{\mathbf{\Gamma}}^{(k)} - \breve{\mathbf{\Gamma}}^{(k),(m)} = \frac{1}{np_{-k}} \sum_{t=1}^{n} \mathbf{X}_{k,t}^{\mathsf{t}} \left( \mathbf{I} - \widehat{\mathbf{D}}_k^{(m)} (\widehat{\mathbf{D}}_k^{(m)})^\top \right) (\mathbf{X}_{k,t}^{\mathsf{t}})^\top,$$

is also non-negative definite, it follows that

$$\frac{1}{p_k} \lambda_j \left( \breve{\mathbf{\Gamma}}^{(k),(m)} \right) \leq \frac{1}{p_k} \lambda_j \left( \widehat{\mathbf{\Gamma}}^{(k)} \right) \leq \begin{cases} \beta_j^{(k)} + O_P(\rho_{n,p}) & \text{for } 1 \leq j \leq r_k, \\ O_P(\rho_{n,p}) & \text{for } j \geq r_k + 1, \end{cases} \tag{B.50}$$

by Weyl's inequality, (B.2) and Assumption 2. From (B.49) and (B.50), we have

$$\frac{\breve{\mu}_j^{(k),(m)}}{\breve{\mu}_{j+1}^{(k),(m)} + \rho_{n,p}} \leq \begin{cases} \frac{\beta_j^{(k)} + O_P(\rho_{n,p})}{\alpha_j^{(k)} + O_P(\rho_{n,p})} = O_P(1) & \text{for } 1 \leq j \leq r_k - 1, \\ \frac{O_P(\rho_{n,p})}{\rho_{n,p}} = O_P(1) & \text{for } j \geq r_k + 1, \end{cases}$$

and

$$\frac{\breve{\mu}_{r_k}^{(k),(m)}}{\breve{\mu}_{r_k+1}^{(k),(m)} + \rho_{n,p}} \geq \frac{\alpha_{r_k}^{(k)} + O_P(\rho_{n,p})}{\rho_{n,p}} \asymp \rho_{n,p}^{-1} \to \infty$$

as $\min(n, p_1, \ldots, p_k) \to \infty$, which completes the proof.

# C Complete simulation results

## C.1 Tensor time series

### C.1.1 Set-up

**Data generation.** The following data generating process is taken from Barigozzi et al. (2023) with modifications to augment the difficulty of the estimation problem. We draw entries of $\mathbf{\Lambda}_k$ independently from Unif$[-1, 1]$. With $\phi = \psi = 0.3$, we introduce serial dependence to $\mathcal{F}_t^\circ$ and $\boldsymbol{\xi}_t^\circ$ as

$$\mathsf{vec}(\mathcal{F}_t^\circ) = \phi \cdot \mathsf{vec}(\mathcal{F}_{t-1}^\circ) + \sqrt{1-\phi^2} \cdot \mathbf{e}_t, \qquad (C.1)$$
$$\mathsf{vec}(\boldsymbol{\xi}_t^\circ) = \psi \cdot \mathsf{vec}(\boldsymbol{\xi}_{t-1}^\circ) + \sqrt{1-\psi^2} \cdot \mathsf{vec}(\mathcal{V}_t).$$

Here, we generate $\mathcal{V}_t \in \mathbb{R}^{p_1 \times p_2 \times p_3}$ such that $\mathsf{vec}(\mathcal{V}_t) = \otimes_{k=K}^1 \mathbf{\Sigma}_k^{1/2} \mathbf{v}_t$, where each $\mathbf{\Sigma}_k$ has $p_k^{-1}$ in the off-diagonals and ones on the diagonal. Also, $\mathbf{e}_t \in \mathbb{R}^r$ and $\mathbf{v}_t \in \mathbb{R}^p$ have i.i.d. zero-mean random elements drawn from either the standard normal distribution and the scaled $t_3$ distribution such that $\mathsf{Var}(e_{jt}) = \mathsf{Var}(v_{it}) = 1$. Note that the cross-sectional correlations induced by the above $\mathbf{\Sigma}_k$ do not conform to the strong mixing condition made in Assumption 5. Nonetheless, the results show that the proposed tail-robust methods are able to handle mild cross-sectional correlations, see also Appendix C.1.5 for the results obtained under the scenarios where $\mathbf{\Sigma}_k$ are Toeptliz matrices.
We consider

(T1) $p_1 = p_2 = p_3 = 10$,

(T2) $p_1 = 100$ and $p_2 = p_3 = 10$,

(T3) $p_1 = 20$, $p_2 = 30$ and $p_3 = 40$,

while varying $n \in \{100, 200, 500\}$. Once $\mathcal{X}_t^\circ = \boldsymbol{\chi}_t^\circ + \boldsymbol{\xi}_t^\circ$ is generated, we additionally consider the situations where the observed data $\mathcal{X}_t$ are contaminated by outliers:

(O1) Outliers are introduced to the idiosyncratic component. Specifically, we randomly select $\mathcal{O} \subset \prod_{k=1}^K [p_k] \times [n]$ with its cardinality $|\mathcal{O}| = [\varrho n p]$. Then for $(\mathbf{i}, t) \in \mathcal{O}$, we set $X_{\mathbf{i},t} = s_{\mathbf{i},t} \cdot U_{\mathbf{i},t}$ with $s_{\mathbf{i},t} \sim_{\text{iid}}$ Unif$\{-1, 1\}$ and $U_{\mathbf{i},t} \sim_{\text{iid}}$ Unif$[Q + 12, Q + 15]$ with $Q$ set to be the $\max(1 - 100/(np), 0.999)$-quantile of $|X_{\mathbf{i},t}^\circ|$; otherwise $X_{\mathbf{i},t} = X_{\mathbf{i},t}^\circ$ if $(\mathbf{i}, t) \notin \mathcal{O}$.

(O2) Outliers are introduced to the factors. Specifically, we randomly select $\mathcal{O} \subset \prod_{k=1}^K [r_k] \times [n]$ with its cardinality $|\mathcal{O}| = [\varrho n r]$. Then for $(\mathbf{j}, t) \in \mathcal{O}$, we set $f_{\mathbf{j},t} = s_{\mathbf{j},t} \cdot U_{\mathbf{j},t}$ with $s_{\mathbf{j},t} \sim_{\text{iid}}$ Unif$\{-1, 1\}$ and $U_{\mathbf{j},t} \sim_{\text{iid}}$ Unif$[Q + 12, Q + 15]$ with $Q$ set to be the $\max(1 - 100/(nr), 0.999)$-quantile of $|f_{\mathbf{j},t}^\circ|$, while $f_{\mathbf{j},t} = f_{\mathbf{j},t}^\circ$ otherwise.

Either under (O1) or (O2), when $\varrho = 0$, there are no outliers and we have $X_{\mathbf{i},t} = X_{\mathbf{i},t}^\circ$ for all $\mathbf{i}$ and $t$.

**Performance assessment.** To assess the performance of any estimator $\widehat{\mathbf{\Lambda}}_k \in \mathbb{R}^{p_k \times r_k}$ in loading space estimation, we compute

$$\mathrm{Err}_{\Lambda_k} = \sqrt{1 - \mathrm{tr}\left(\Pi_{\widehat{\mathbf{\Lambda}}_k} \Pi_{\mathbf{\Lambda}_k}\right)/r_k}, \quad \text{where} \quad \Pi_{\mathbf{A}} = \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1}\mathbf{A}^\top. \tag{C.2}$$

To assess the quality in common component estimation, for any estimator $\widehat{\boldsymbol{\chi}}_t$, we evaluate

$$\mathrm{Err}_\chi(\mathcal{T}) = \frac{\sum_{t \in \mathcal{T}} |\widehat{\boldsymbol{\chi}}_t - \boldsymbol{\chi}_t|_2^2}{\sum_{t \in \mathcal{T}} |\boldsymbol{\chi}_t|_2^2} \tag{C.3}$$

with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local').

### C.1.2 Estimation of loadings and common component

**No outlier.** See Figures C.1–C.2 and Tables C.1–C.2 for the results from the loading and the common component estimation obtained under (T1)–(T3), in the absence of any outlier.

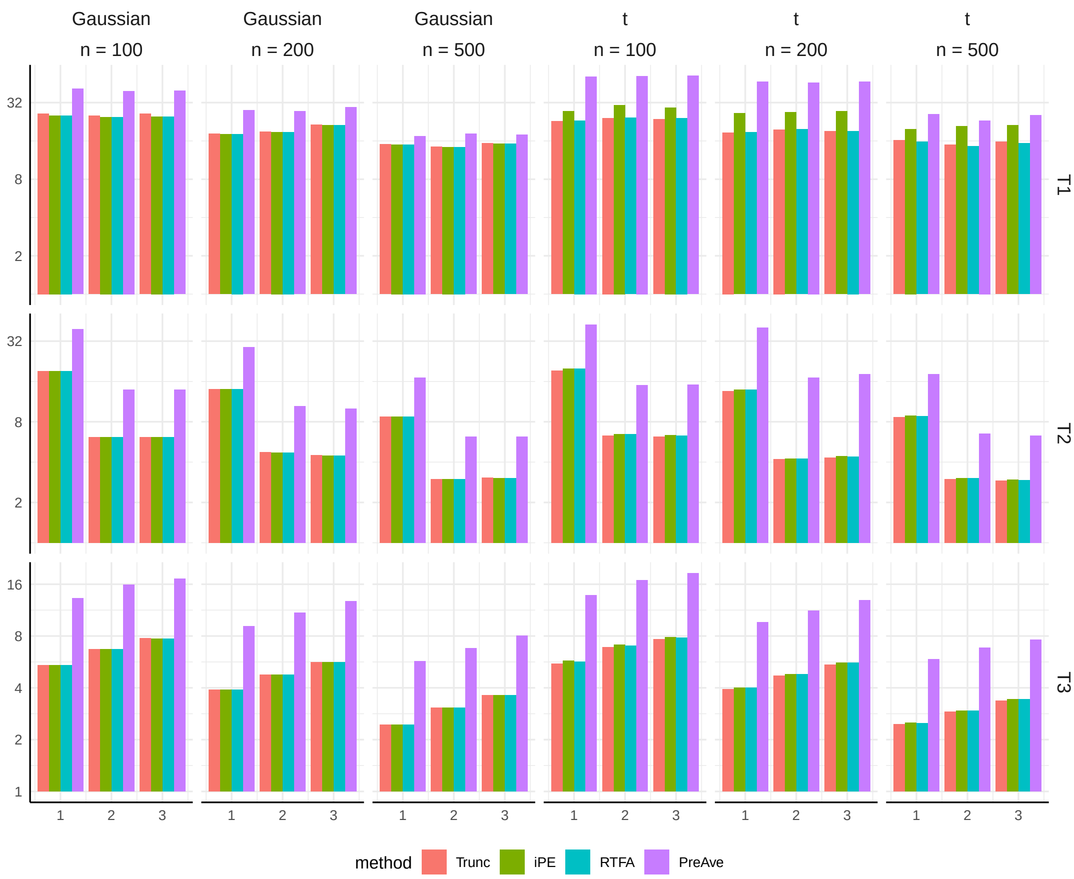


Figure C.1: Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$ and the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) in the absence of any outlier, averaged over 100 realisations per setting, for (T1)–(T3) (top to bottom). In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

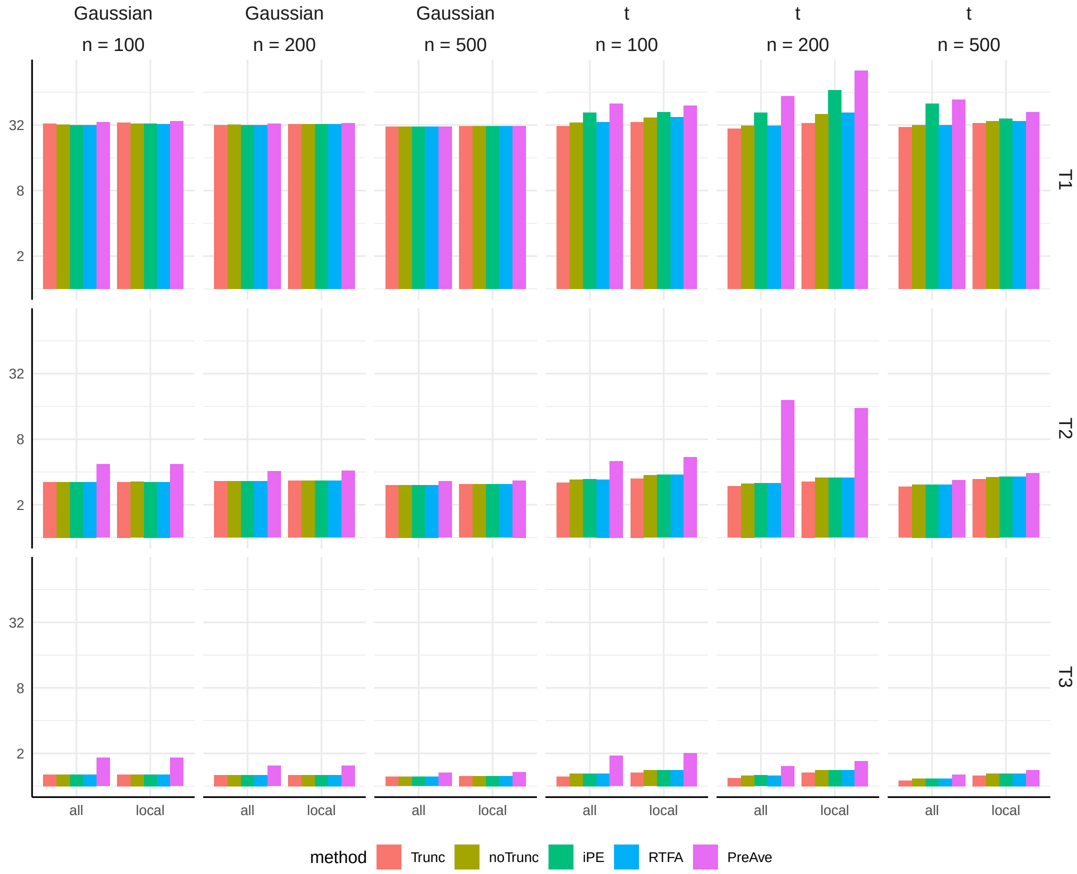


Figure C.2: Common component estimation errors measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') for Trunc, noTrunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$ and the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) in the absence of any outlier, averaged over 100 realisations per setting, for (T1)–(T3) (top to bottom). In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

Table C.1: Loading estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.2) for each mode scaled by 100, over varying $n \in \{100, 200, 500\}$ and the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) in the absence of any outlier. We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | $n$ | Dist | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| (T1) | 100 | Gaussian | 1 | 2.628 | 1.723 | 2.536 | 1.209 | 2.537 | 1.211 | 4.122 | 1.678 |
| | | | 2 | 2.528 | 1.305 | 2.469 | 1.216 | 2.469 | 1.214 | 3.953 | 1.168 |
| | | | 3 | 2.616 | 1.908 | 2.488 | 1.15 | 2.487 | 1.145 | 3.976 | 1.387 |
| | | $t_3$ | 1 | 2.288 | 1.018 | 2.733 | 3.586 | 2.319 | 1.036 | 5.139 | 5.921 |
| | | | 2 | 2.425 | 1.045 | 3.065 | 5.603 | 2.438 | 1.042 | 5.189 | 6.777 |
| | | | 3 | 2.368 | 0.811 | 2.924 | 4.276 | 2.421 | 0.864 | 5.199 | 5.565 |
| | 200 | Gaussian | 1 | 1.821 | 0.852 | 1.81 | 0.839 | 1.81 | 0.841 | 2.79 | 0.913 |
| | | | 2 | 1.892 | 0.789 | 1.879 | 0.778 | 1.879 | 0.779 | 2.741 | 0.839 |
| | | | 3 | 2.141 | 1.653 | 2.124 | 1.623 | 2.124 | 1.626 | 2.94 | 1.347 |
| | | $t_3$ | 1 | 1.86 | 0.819 | 2.637 | 4.933 | 1.874 | 0.844 | 4.672 | 10.607 |
| | | | 2 | 1.969 | 0.947 | 2.692 | 4.92 | 1.973 | 1.013 | 4.604 | 9.930 |
| | | | 3 | 1.909 | 0.75 | 2.743 | 5.998 | 1.916 | 0.753 | 4.67 | 8.578 |
| | 500 | Gaussian | 1 | 1.504 | 0.769 | 1.497 | 0.76 | 1.497 | 0.76 | 1.751 | 0.600 |
| | | | 2 | 1.439 | 0.658 | 1.433 | 0.65 | 1.433 | 0.65 | 1.825 | 0.697 |
| | | | 3 | 1.53 | 0.928 | 1.523 | 0.919 | 1.523 | 0.92 | 1.789 | 0.598 |
| | | $t_3$ | 1 | 1.619 | 0.937 | 1.984 | 3.444 | 1.576 | 0.899 | 2.594 | 7.028 |
| | | | 2 | 1.496 | 0.901 | 2.091 | 5.507 | 1.451 | 0.855 | 2.315 | 5.518 |
| | | | 3 | 1.574 | 1.146 | 2.124 | 5.016 | 1.535 | 1.124 | 2.548 | 6.563 |
| (T2) | 100 | Gaussian | 1 | 1.91 | 0.296 | 1.91 | 0.297 | 1.91 | 0.297 | 3.948 | 0.882 |
| | | | 2 | 0.618 | 0.184 | 0.616 | 0.182 | 0.616 | 0.182 | 1.398 | 0.472 |
| | | | 3 | 0.616 | 0.172 | 0.613 | 0.17 | 0.613 | 0.17 | 1.398 | 0.475 |
| | | $t_3$ | 1 | 1.932 | 0.302 | 2 | 0.338 | 1.991 | 0.334 | 4.263 | 0.990 |
| | | | 2 | 0.634 | 0.192 | 0.649 | 0.207 | 0.647 | 0.205 | 1.501 | 0.533 |
| | | | 3 | 0.623 | 0.198 | 0.636 | 0.198 | 0.633 | 0.196 | 1.514 | 0.524 |
| | 200 | Gaussian | 1 | 1.404 | 0.183 | 1.404 | 0.183 | 1.404 | 0.183 | 2.883 | 0.617 |
| | | | 2 | 0.474 | 0.144 | 0.471 | 0.143 | 0.472 | 0.143 | 1.048 | 0.458 |
| | | | 3 | 0.452 | 0.121 | 0.45 | 0.119 | 0.45 | 0.119 | 1.006 | 0.285 |
| | | $t_3$ | 1 | 1.358 | 0.179 | 1.398 | 0.19 | 1.39 | 0.188 | 4.061 | 7.863 |
| | | | 2 | 0.421 | 0.106 | 0.427 | 0.106 | 0.424 | 0.106 | 1.705 | 6.855 |
| | | | 3 | 0.434 | 0.117 | 0.442 | 0.123 | 0.439 | 0.121 | 1.816 | 7.468 |
| | 500 | Gaussian | 1 | 0.873 | 0.123 | 0.872 | 0.123 | 0.872 | 0.123 | 1.711 | 0.341 |
| | | | 2 | 0.3 | 0.105 | 0.299 | 0.105 | 0.299 | 0.105 | 0.62 | 0.189 |
| | | | 3 | 0.306 | 0.104 | 0.305 | 0.103 | 0.305 | 0.103 | 0.62 | 0.169 |
| | | $t_3$ | 1 | 0.87 | 0.125 | 0.889 | 0.128 | 0.885 | 0.129 | 1.815 | 0.388 |
| | | | 2 | 0.298 | 0.111 | 0.304 | 0.114 | 0.303 | 0.113 | 0.655 | 0.207 |
| | | | 3 | 0.293 | 0.093 | 0.296 | 0.093 | 0.294 | 0.094 | 0.632 | 0.245 |
| (T3) | 100 | Gaussian | 1 | 0.541 | 0.09 | 0.541 | 0.091 | 0.541 | 0.091 | 1.329 | 0.276 |
| | | | 2 | 0.671 | 0.093 | 0.671 | 0.092 | 0.671 | 0.092 | 1.591 | 0.276 |
| | | | 3 | 0.775 | 0.093 | 0.774 | 0.093 | 0.774 | 0.093 | 1.719 | 0.260 |
| | | $t_3$ | 1 | 0.554 | 0.087 | 0.574 | 0.096 | 0.569 | 0.092 | 1.381 | 0.288 |
| | | | 2 | 0.69 | 0.108 | 0.71 | 0.11 | 0.704 | 0.102 | 1.69 | 0.374 |
| | | | 3 | 0.764 | 0.095 | 0.787 | 0.102 | 0.783 | 0.099 | 1.856 | 0.329 |
| | 200 | Gaussian | 1 | 0.39 | 0.066 | 0.39 | 0.066 | 0.39 | 0.066 | 0.913 | 0.155 |
| | | | 2 | 0.478 | 0.063 | 0.478 | 0.063 | 0.478 | 0.063 | 1.092 | 0.163 |
| | | | 3 | 0.563 | 0.069 | 0.563 | 0.069 | 0.563 | 0.069 | 1.278 | 0.237 |
| | | $t_3$ | 1 | 0.392 | 0.065 | 0.401 | 0.068 | 0.4 | 0.067 | 0.963 | 0.221 |
| | | | 2 | 0.47 | 0.07 | 0.48 | 0.074 | 0.479 | 0.074 | 1.124 | 0.180 |
| | | | 3 | 0.547 | 0.073 | 0.56 | 0.077 | 0.558 | 0.076 | 1.292 | 0.225 |
| | 500 | Gaussian | 1 | 0.245 | 0.032 | 0.245 | 0.032 | 0.245 | 0.032 | 0.573 | 0.092 |
| | | | 2 | 0.306 | 0.039 | 0.306 | 0.039 | 0.306 | 0.039 | 0.682 | 0.094 |
| | | | 3 | 0.362 | 0.043 | 0.361 | 0.043 | 0.361 | 0.043 | 0.801 | 0.104 |
| | | $t_3$ | 1 | 0.246 | 0.036 | 0.251 | 0.036 | 0.25 | 0.036 | 0.586 | 0.110 |
| | | | 2 | 0.29 | 0.032 | 0.295 | 0.033 | 0.294 | 0.033 | 0.686 | 0.097 |
| | | | 3 | 0.338 | 0.039 | 0.344 | 0.04 | 0.343 | 0.04 | 0.763 | 0.110 |

Table C.2: Common component estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.2) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') scaled by 1000, over varying $n \in \{100, 200, 500\}$ and the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | noTrunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | $n$ | Dist | Range | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| (T1) | 100 | Gaussian | All | 33.059 | 15.354 | 32.308 | 10.698 | 32.055 | 10.013 | 32.054 | 10.011 | 34.086 | 10.340 |
| | | | Local | 33.672 | 15.405 | 33.029 | 11.869 | 32.774 | 11.258 | 32.773 | 11.256 | 34.76 | 11.408 |
| | | $t_3$ | All | 31.183 | 13.386 | 33.805 | 15.601 | 41.629 | 78.458 | 33.892 | 15.701 | 50.089 | 120.027 |
| | | | Local | 33.997 | 18.401 | 37.496 | 25.511 | 41.886 | 47.967 | 37.619 | 25.818 | 47.986 | 66.670 |
| | 200 | Gaussian | All | 32.088 | 12.048 | 32.097 | 12.054 | 32.076 | 12.021 | 32.076 | 12.022 | 32.781 | 11.908 |
| | | | Local | 32.546 | 13.063 | 32.553 | 13.08 | 32.532 | 13.049 | 32.533 | 13.051 | 33.24 | 12.926 |
| | | $t_3$ | All | 29.567 | 8.576 | 31.541 | 10.577 | 41.377 | 80.622 | 31.617 | 10.966 | 58.644 | 161.159 |
| | | | Local | 33.361 | 14.938 | 40.306 | 58.31 | 67.13 | 291.068 | 41.535 | 70.138 | 101.211 | 522.470 |
| | 500 | Gaussian | All | 30.972 | 9.427 | 30.979 | 9.43 | 30.971 | 9.42 | 30.971 | 9.421 | 30.965 | 9.211 |
| | | | Local | 31.368 | 9.513 | 31.374 | 9.514 | 31.367 | 9.509 | 31.367 | 9.509 | 31.376 | 9.441 |
| | | $t_3$ | All | 30.688 | 11.532 | 31.927 | 13.292 | 50.174 | 188.124 | 31.918 | 13.429 | 54.479 | 233.345 |
| | | | Local | 33.396 | 15.469 | 34.785 | 16.551 | 36.763 | 24.055 | 34.744 | 16.527 | 42.108 | 74.449 |
| (T2) | 100 | Gaussian | All | 3.239 | 0.968 | 3.239 | 0.968 | 3.238 | 0.967 | 3.238 | 0.967 | 4.703 | 1.638 |
| | | | Local | 3.239 | 1.062 | 3.24 | 1.063 | 3.239 | 1.062 | 3.239 | 1.062 | 4.703 | 1.693 |
| | | $t_3$ | All | 3.189 | 0.958 | 3.385 | 1.061 | 3.416 | 1.081 | 3.412 | 1.079 | 5.048 | 1.759 |
| | | | Local | 3.488 | 1.246 | 3.73 | 1.463 | 3.769 | 1.511 | 3.76 | 1.488 | 5.46 | 2.130 |
| | 200 | Gaussian | All | 3.295 | 0.86 | 3.295 | 0.86 | 3.295 | 0.859 | 3.295 | 0.859 | 4.07 | 1.154 |
| | | | Local | 3.323 | 0.996 | 3.323 | 0.996 | 3.323 | 0.996 | 3.323 | 0.996 | 4.102 | 1.274 |
| | | $t_3$ | All | 2.959 | 0.733 | 3.133 | 0.811 | 3.146 | 0.817 | 3.143 | 0.816 | 18.284 | 118.031 |
| | | | Local | 3.273 | 1.189 | 3.54 | 1.459 | 3.556 | 1.476 | 3.552 | 1.47 | 15.347 | 86.214 |
| | 500 | Gaussian | All | 3.037 | 0.867 | 3.037 | 0.868 | 3.037 | 0.867 | 3.037 | 0.867 | 3.3 | 0.954 |
| | | | Local | 3.077 | 0.976 | 3.077 | 0.976 | 3.077 | 0.976 | 3.077 | 0.976 | 3.34 | 1.062 |
| | | $t_3$ | All | 2.932 | 0.863 | 3.068 | 0.917 | 3.073 | 0.919 | 3.072 | 0.919 | 3.371 | 1.025 |
| | | | Local | 3.421 | 1.267 | 3.599 | 1.485 | 3.604 | 1.489 | 3.602 | 1.486 | 3.91 | 1.577 |
| (T3) | 100 | Gaussian | All | 1.28 | 0.252 | 1.28 | 0.251 | 1.28 | 0.251 | 1.28 | 0.251 | 1.824 | 0.385 |
| | | | Local | 1.281 | 0.29 | 1.281 | 0.291 | 1.281 | 0.291 | 1.281 | 0.291 | 1.825 | 0.416 |
| | | $t_3$ | All | 1.227 | 0.267 | 1.304 | 0.305 | 1.313 | 0.311 | 1.31 | 0.305 | 1.905 | 0.485 |
| | | | Local | 1.334 | 0.411 | 1.403 | 0.432 | 1.41 | 0.433 | 1.408 | 0.433 | 2.024 | 0.587 |
| | 200 | Gaussian | All | 1.265 | 0.269 | 1.265 | 0.269 | 1.265 | 0.269 | 1.265 | 0.269 | 1.542 | 0.342 |
| | | | Local | 1.272 | 0.333 | 1.272 | 0.333 | 1.272 | 0.333 | 1.272 | 0.333 | 1.552 | 0.399 |
| | | $t_3$ | All | 1.194 | 0.245 | 1.257 | 0.269 | 1.26 | 0.271 | 1.259 | 0.271 | 1.538 | 0.330 |
| | | | Local | 1.339 | 0.433 | 1.406 | 0.45 | 1.409 | 0.45 | 1.409 | 0.45 | 1.698 | 0.510 |
| | 500 | Gaussian | All | 1.222 | 0.215 | 1.222 | 0.215 | 1.222 | 0.215 | 1.222 | 0.215 | 1.328 | 0.231 |
| | | | Local | 1.246 | 0.28 | 1.246 | 0.28 | 1.246 | 0.28 | 1.246 | 0.28 | 1.353 | 0.295 |
| | | $t_3$ | All | 1.13 | 0.205 | 1.178 | 0.218 | 1.179 | 0.218 | 1.178 | 0.218 | 1.279 | 0.237 |
| | | | Local | 1.252 | 0.459 | 1.303 | 0.49 | 1.304 | 0.49 | 1.304 | 0.49 | 1.408 | 0.504 |

**Outliers in the idiosyncratic component.** See Figures C.3–C.8 and Tables C.3–C.8 for the results from loading and common component estimation obtained under (T1)–(T3) with outliers in the idiosyncratic component under (O1).

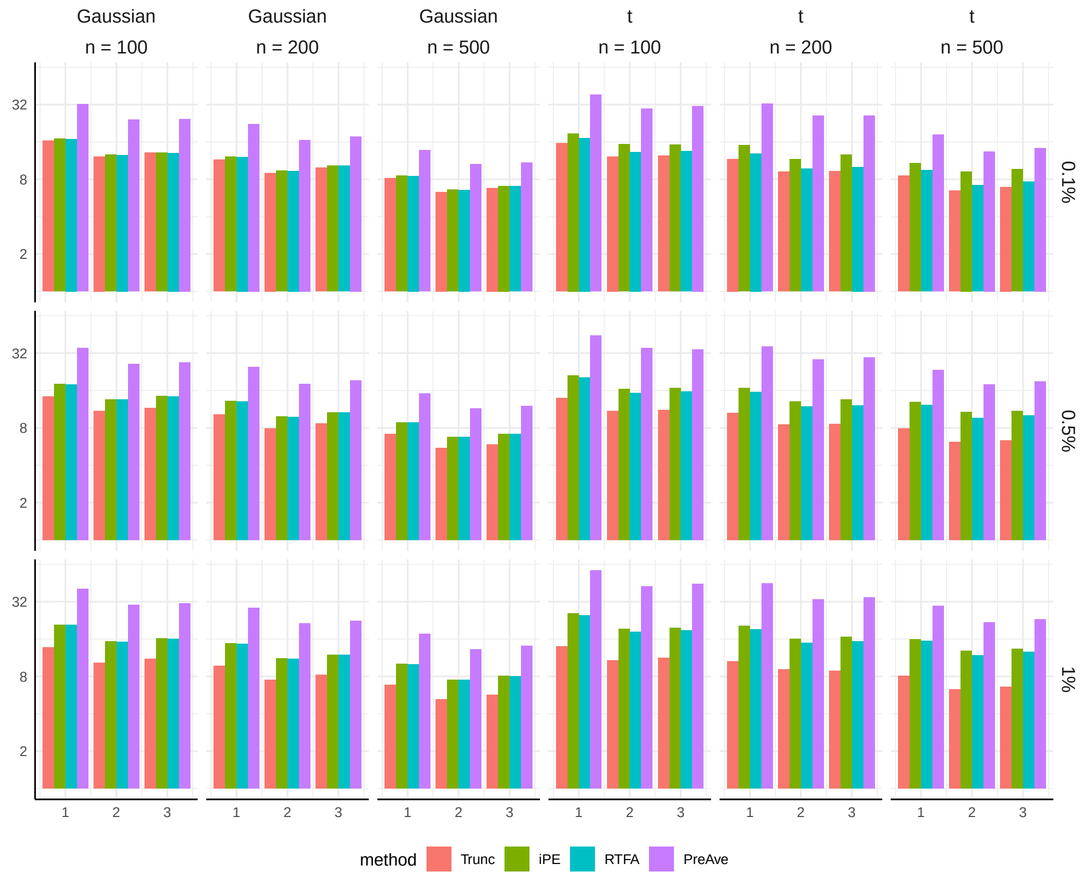


Figure C.3: (T1) Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

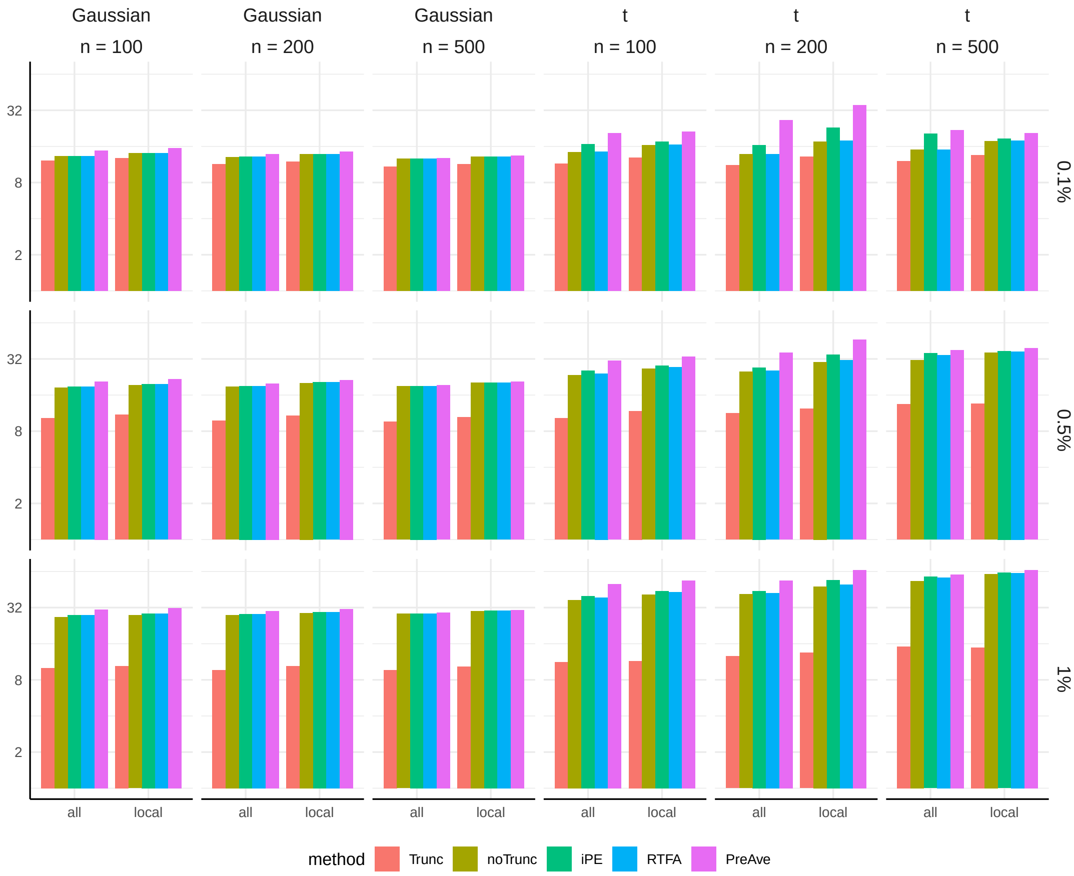


Figure C.4: (T1) Common component estimation errors measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n-10+1, \ldots, n\}$ ('local') for Trunc, noTrunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

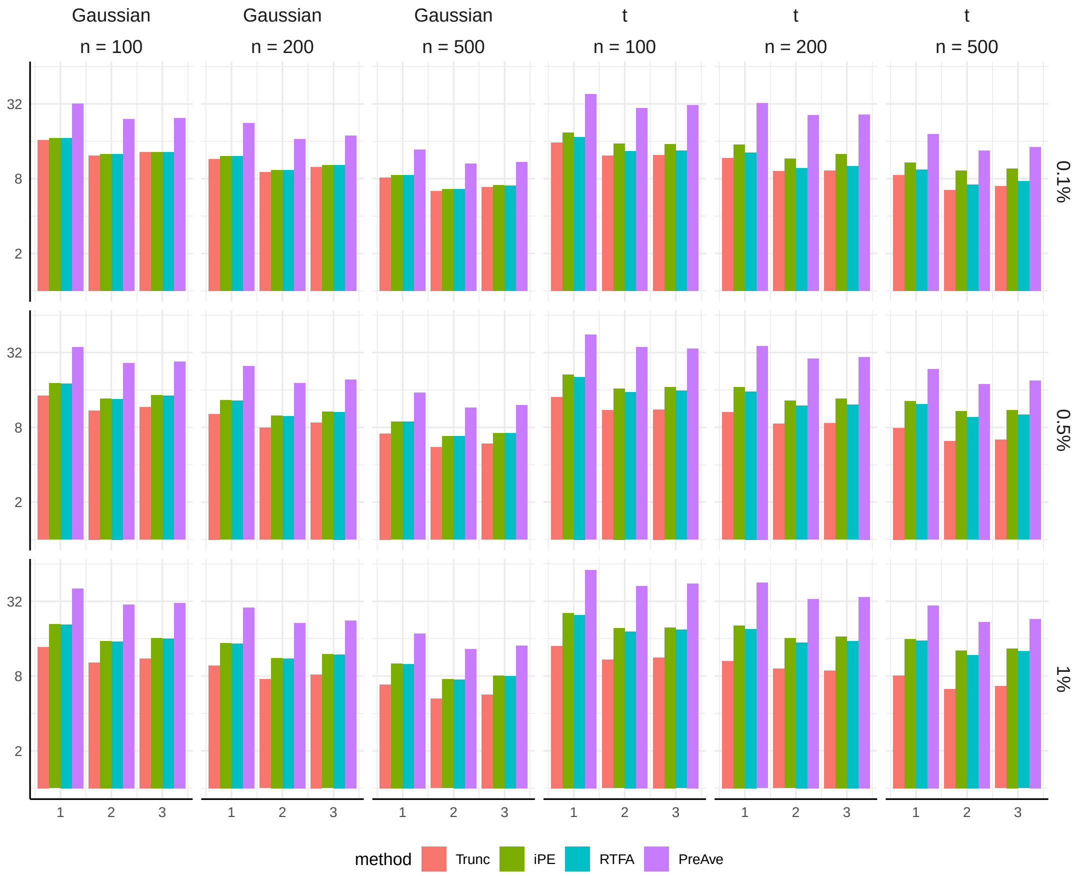


Figure C.5: (T2) Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

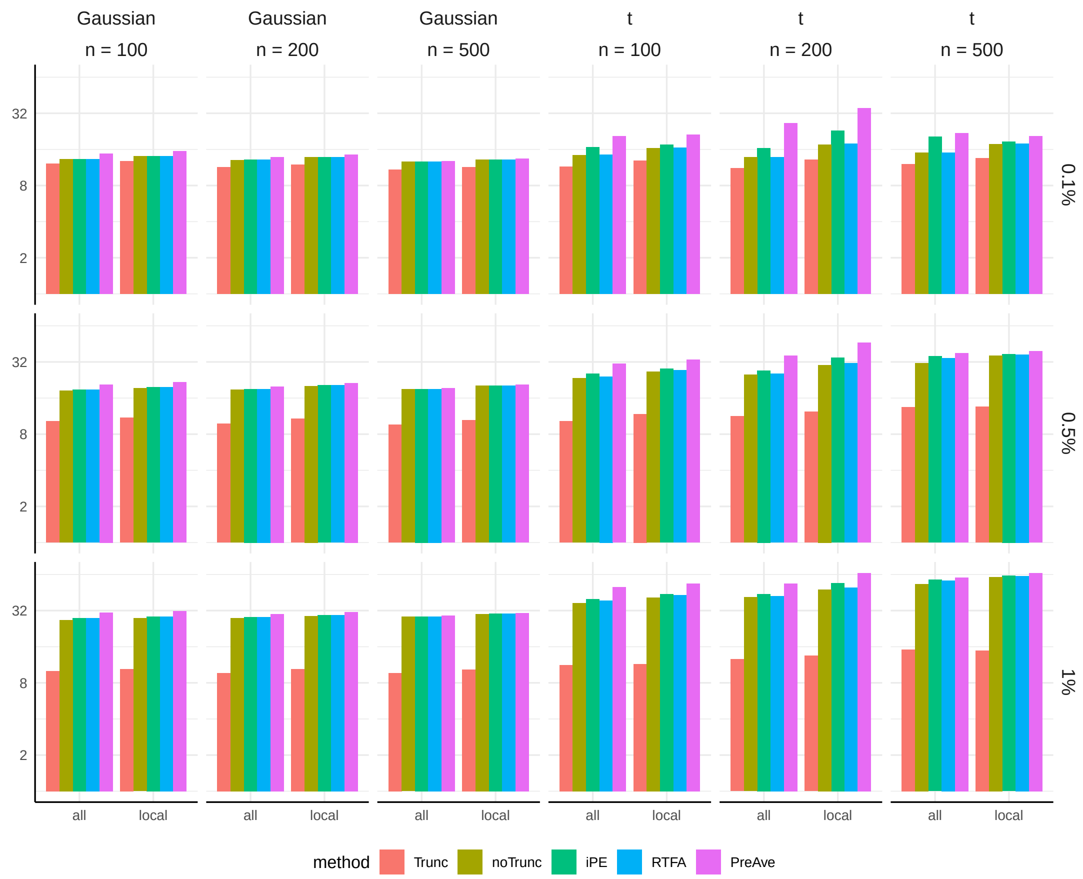


Figure C.6: (T2) Common component estimation errors measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n-10+1, \ldots, n\}$ ('local') for Trunc, noTrunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

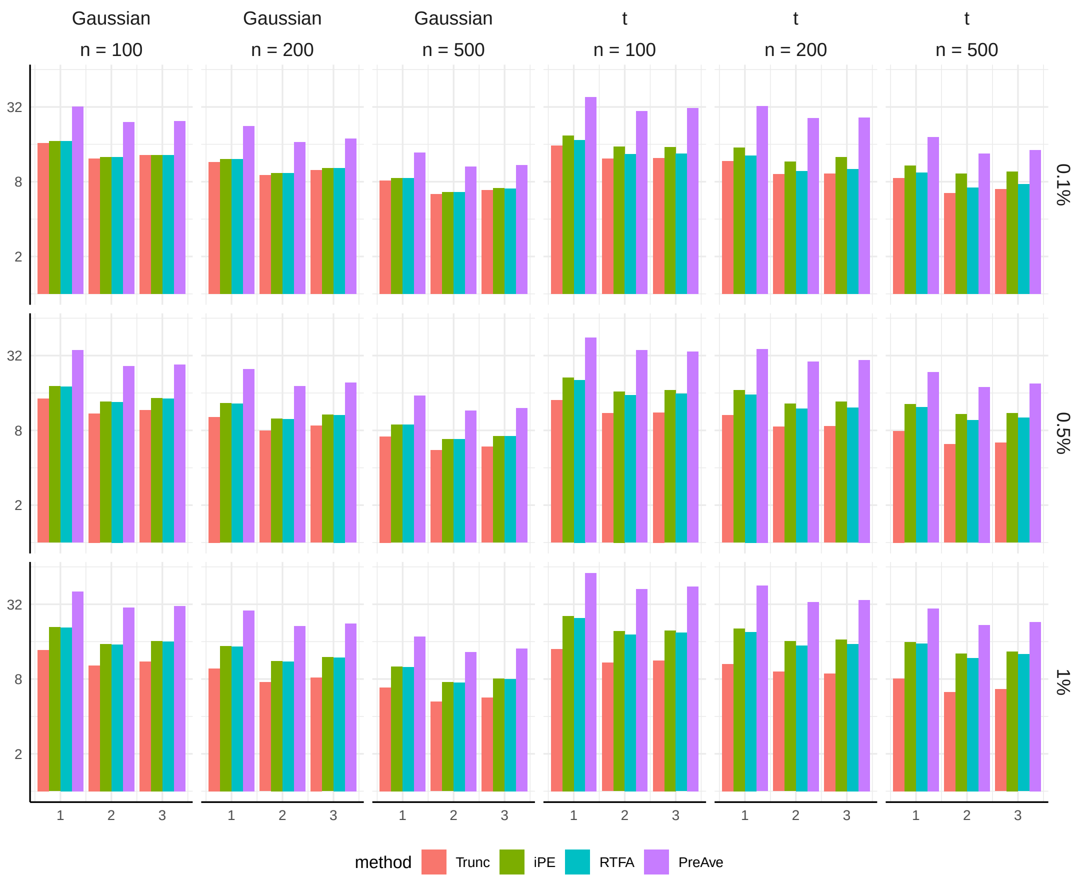


Figure C.7: (T3) Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

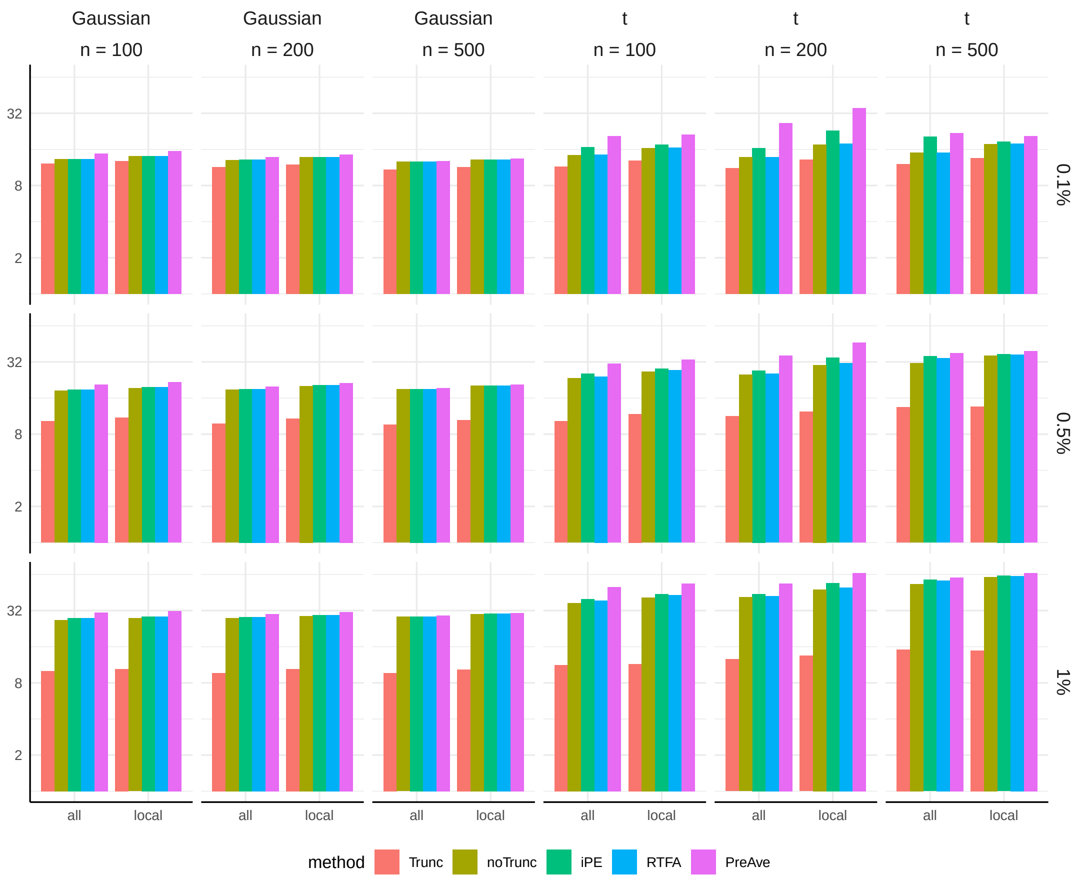


Figure C.8: (T3) Common component estimation errors measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n-10+1, \ldots, n\}$ ('local') for Trunc, noTrunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

Table C.3: (T1) Loading estimation errors of Trunc, iPE, RTFA and PreAve measured as in (C.2) for each mode scaled by 100, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | 1 | 2.627 | 1.391 | 2.827 | 1.382 | 2.803 | 1.358 | 4.656 | 1.872 |
| | | | 2 | 2.515 | 1.298 | 2.694 | 1.29 | 2.669 | 1.27 | 4.417 | 1.274 |
| | | | 3 | 2.619 | 1.651 | 2.799 | 1.576 | 2.77 | 1.532 | 4.426 | 1.421 |
| | | 0.5 | 1 | 2.786 | 1.834 | 3.718 | 1.426 | 3.663 | 1.41 | 6.316 | 2.185 |
| | | | 2 | 2.717 | 1.391 | 3.709 | 1.434 | 3.654 | 1.413 | 6.173 | 1.471 |
| | | | 3 | 3.057 | 4.338 | 4.125 | 4.559 | 4.025 | 4.117 | 6.316 | 2.449 |
| | | 1 | 1 | 2.955 | 1.611 | 4.985 | 2.104 | 4.922 | 2.046 | 8.471 | 3.368 |
| | | | 2 | 2.796 | 1.354 | 4.641 | 1.729 | 4.577 | 1.703 | 8.111 | 2.426 |
| | | | 3 | 3.146 | 3.55 | 5.012 | 3.651 | 4.931 | 3.557 | 8.23 | 3.057 |
| | $t_3$ | 0.1 | 1 | 2.445 | 1.13 | 3.18 | 3.723 | 2.724 | 1.137 | 6.257 | 6.588 |
| | | | 2 | 2.532 | 1.046 | 3.475 | 5.56 | 2.84 | 1.071 | 5.865 | 6.362 |
| | | | 3 | 2.492 | 0.811 | 3.304 | 4.258 | 2.791 | 0.857 | 6.574 | 7.600 |
| | | 0.5 | 1 | 2.704 | 1.117 | 4.467 | 3.796 | 4 | 1.403 | 8.454 | 7.001 |
| | | | 2 | 2.853 | 1.29 | 4.781 | 5.555 | 4.178 | 1.584 | 8.909 | 8.310 |
| | | | 3 | 2.822 | 1.014 | 4.787 | 4.504 | 4.251 | 1.726 | 8.275 | 6.292 |
| | | 1 | 1 | 3.07 | 1.5 | 5.939 | 4.143 | 5.491 | 2.282 | 11.769 | 9.398 |
| | | | 2 | 3.064 | 1.276 | 6.105 | 5.561 | 5.479 | 1.935 | 11.781 | 7.872 |
| | | | 3 | 3.174 | 1.221 | 5.925 | 4.357 | 5.45 | 1.821 | 12.107 | 8.759 |
| 200 | Gaussian | 0.1 | 1 | 1.867 | 0.879 | 2.033 | 0.924 | 2.014 | 0.925 | 3.157 | 1.030 |
| | | | 2 | 1.917 | 0.803 | 2.052 | 0.828 | 2.037 | 0.829 | 3.087 | 0.895 |
| | | | 3 | 2.176 | 1.716 | 2.304 | 1.61 | 2.293 | 1.623 | 3.303 | 1.474 |
| | | 0.5 | 1 | 2.001 | 1.087 | 2.763 | 1.386 | 2.72 | 1.378 | 4.52 | 1.341 |
| | | | 2 | 2.02 | 0.817 | 2.699 | 1.027 | 2.67 | 1.013 | 4.28 | 1.242 |
| | | | 3 | 2.276 | 1.822 | 2.943 | 2.068 | 2.906 | 2.01 | 4.456 | 1.793 |
| | | 1 | 1 | 2.028 | 0.894 | 3.314 | 1.379 | 3.278 | 1.357 | 5.835 | 1.885 |
| | | | 2 | 2.114 | 0.818 | 3.454 | 1.186 | 3.417 | 1.174 | 5.874 | 1.824 |
| | | | 3 | 2.356 | 1.905 | 3.67 | 2.54 | 3.638 | 2.494 | 5.941 | 2.680 |
| | $t_3$ | 0.1 | 1 | 2.059 | 1.179 | 2.975 | 4.911 | 2.224 | 1.011 | 5.241 | 10.557 |
| | | | 2 | 2.181 | 1.781 | 2.955 | 4.812 | 2.24 | 0.969 | 5.217 | 10.078 |
| | | | 3 | 2.128 | 1.11 | 3.083 | 5.96 | 2.243 | 0.861 | 5.318 | 8.649 |
| | | 0.5 | 1 | 2.261 | 1.009 | 3.811 | 4.923 | 3.097 | 1.229 | 7.055 | 11.282 |
| | | | 2 | 2.475 | 1.76 | 4.13 | 5.13 | 3.459 | 2.118 | 7.449 | 11.591 |
| | | | 3 | 2.425 | 1.194 | 4.143 | 5.903 | 3.345 | 1.417 | 7.345 | 9.450 |
| | | 1 | 1 | 2.428 | 0.984 | 4.839 | 4.977 | 4.15 | 1.606 | 8.912 | 10.954 |
| | | | 2 | 2.951 | 3.819 | 5.378 | 6.324 | 4.685 | 4.441 | 9.095 | 11.277 |
| | | | 3 | 2.634 | 1.539 | 5.365 | 6.085 | 4.565 | 1.867 | 9.091 | 9.797 |
| 500 | Gaussian | 0.1 | 1 | 1.517 | 0.767 | 1.588 | 0.756 | 1.576 | 0.757 | 2.044 | 0.697 |
| | | | 2 | 1.455 | 0.667 | 1.554 | 0.693 | 1.542 | 0.693 | 2.068 | 0.710 |
| | | | 3 | 1.544 | 0.942 | 1.62 | 0.957 | 1.61 | 0.959 | 2.055 | 0.659 |
| | | 0.5 | 1 | 1.562 | 0.762 | 1.981 | 0.82 | 1.955 | 0.811 | 2.842 | 0.827 |
| | | | 2 | 1.504 | 0.658 | 1.951 | 0.776 | 1.93 | 0.762 | 2.765 | 0.810 |
| | | | 3 | 1.598 | 0.92 | 2.027 | 0.987 | 2.005 | 0.97 | 2.861 | 0.891 |
| | | 1 | 1 | 1.66 | 0.875 | 2.48 | 1.033 | 2.46 | 1.023 | 3.655 | 1.212 |
| | | | 2 | 1.579 | 0.684 | 2.42 | 0.928 | 2.396 | 0.927 | 3.551 | 1.206 |
| | | | 3 | 1.687 | 0.91 | 2.514 | 1.088 | 2.482 | 1.082 | 3.736 | 0.997 |
| | $t_3$ | 0.1 | 1 | 1.711 | 0.982 | 2.19 | 3.427 | 1.785 | 0.936 | 3.043 | 7.266 |
| | | | 2 | 1.585 | 0.904 | 2.375 | 5.595 | 1.749 | 1.493 | 2.782 | 5.557 |
| | | | 3 | 1.679 | 1.216 | 2.462 | 5.175 | 1.843 | 1.571 | 3.045 | 7.105 |
| | | 0.5 | 1 | 1.875 | 0.934 | 2.928 | 3.469 | 2.827 | 3.296 | 4.46 | 7.377 |
| | | | 2 | 1.823 | 1.043 | 3.178 | 5.649 | 3.057 | 5.625 | 4.391 | 7.050 |
| | | | 3 | 1.855 | 1.401 | 3.284 | 5.537 | 3.167 | 5.334 | 4.795 | 8.137 |
| | | 1 | 1 | 2.06 | 1.082 | 3.811 | 3.677 | 3.73 | 3.579 | 5.948 | 7.491 |
| | | | 2 | 1.958 | 1.034 | 3.937 | 5.828 | 3.846 | 5.846 | 5.741 | 7.186 |
| | | | 3 | 2.063 | 1.351 | 4.336 | 5.745 | 4.233 | 5.614 | 6.323 | 8.622 |

Table C.4: (T1) Common component estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') scaled by 1000, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | noTrunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Range | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | All | 33.455 | 11.008 | 41.982 | 12.705 | 42.268 | 12.783 | 42.231 | 12.762 | 44.875 | 13.268 |
| | | | Local | 34.236 | 12.184 | 43.343 | 14.76 | 43.629 | 14.814 | 43.59 | 14.804 | 46.221 | 15.120 |
| | | 0.5 | All | 39.472 | 28.619 | 81.07 | 22.955 | 82.76 | 23.121 | 82.629 | 23.004 | 87.389 | 23.500 |
| | | | Local | 40.134 | 28.032 | 83.248 | 26.692 | 84.784 | 26.419 | 84.643 | 26.272 | 89.314 | 26.639 |
| | | 1 | All | 43.663 | 22.63 | 129.657 | 34.885 | 133.993 | 36.913 | 133.83 | 36.82 | 142.188 | 38.378 |
| | | | Local | 44.045 | 23.667 | 133.009 | 40.867 | 137.319 | 42.679 | 137.169 | 42.612 | 145.811 | 44.147 |
| | $t_3$ | 0.1 | All | 34.229 | 14.747 | 47.738 | 19.901 | 56.128 | 80.231 | 48.256 | 20.241 | 68.019 | 123.378 |
| | | | Local | 36.736 | 19.468 | 52.354 | 29.875 | 57.33 | 50.403 | 52.955 | 30.455 | 68.558 | 88.547 |
| | | 0.5 | All | 42.013 | 20.211 | 103.713 | 38.781 | 114.855 | 87.771 | 106.823 | 40.888 | 131.397 | 133.746 |
| | | | Local | 44.458 | 22.906 | 113.305 | 49.078 | 121.201 | 62.55 | 116.917 | 51.737 | 136.548 | 90.990 |
| | | 1 | All | 51.305 | 28.153 | 171.686 | 60.751 | 186.443 | 99.973 | 178.438 | 64.033 | 216.495 | 159.991 |
| | | | Local | 50.076 | 35.75 | 189.954 | 90.185 | 201.948 | 99.06 | 197.386 | 93.25 | 231.41 | 130.627 |
| 200 | Gaussian | 0.1 | All | 33.446 | 12.403 | 42.314 | 14.657 | 42.462 | 14.629 | 42.45 | 14.632 | 43.417 | 14.457 |
| | | | Local | 34.026 | 13.991 | 43.589 | 16.943 | 43.732 | 16.928 | 43.722 | 16.937 | 44.685 | 16.797 |
| | | 0.5 | All | 37.051 | 12.683 | 83.436 | 25.433 | 84.441 | 25.899 | 84.382 | 25.846 | 86.53 | 25.590 |
| | | | Local | 38.104 | 15.155 | 86.765 | 33.922 | 87.772 | 34.597 | 87.72 | 34.54 | 89.781 | 34.020 |
| | | 1 | All | 41.301 | 14.369 | 134.554 | 39.238 | 136.884 | 40.526 | 136.812 | 40.472 | 141 | 41.095 |
| | | | Local | 42.153 | 16.987 | 137.415 | 50.171 | 139.829 | 51.761 | 139.755 | 51.687 | 144.087 | 51.772 |
| | $t_3$ | 0.1 | All | 36.882 | 29.685 | 47.316 | 13.574 | 56.977 | 81.289 | 47.332 | 13.821 | 75.979 | 167.662 |
| | | | Local | 39.758 | 24.284 | 58.658 | 62.066 | 84.756 | 283.443 | 60.097 | 75.538 | 122.24 | 530.052 |
| | | 0.5 | All | 48.936 | 33.266 | 109.872 | 27.741 | 120.934 | 88.181 | 111.324 | 29.002 | 145.495 | 191.421 |
| | | | Local | 47.946 | 28.949 | 129.931 | 85.113 | 156.03 | 279.577 | 134.074 | 106.16 | 199.005 | 535.579 |
| | | 1 | All | 58.339 | 33.636 | 187.416 | 45.39 | 201.597 | 98.858 | 191.49 | 47.834 | 226.3 | 185.902 |
| | | | Local | 59.732 | 61.851 | 215.859 | 115.161 | 252.076 | 353.252 | 224.423 | 145.083 | 292.548 | 571.622 |
| 500 | Gaussian | 0.1 | All | 32.224 | 9.539 | 41.677 | 11.489 | 41.748 | 11.518 | 41.74 | 11.518 | 41.879 | 11.299 |
| | | | Local | 32.563 | 9.652 | 41.692 | 12.974 | 41.763 | 13.007 | 41.755 | 13.006 | 41.873 | 12.870 |
| | | 0.5 | All | 36.617 | 10.658 | 84.782 | 20.29 | 85.231 | 20.485 | 85.208 | 20.469 | 85.817 | 20.326 |
| | | | Local | 37.521 | 12.05 | 88.152 | 26.985 | 88.615 | 27.183 | 88.593 | 27.177 | 89.182 | 27.014 |
| | | 1 | All | 42.041 | 13.695 | 138.582 | 31.401 | 139.616 | 31.787 | 139.582 | 31.779 | 140.812 | 31.727 |
| | | | Local | 43.236 | 13.997 | 144.556 | 37.509 | 145.626 | 38 | 145.595 | 37.996 | 146.83 | 38.043 |
| | $t_3$ | 0.1 | All | 42.516 | 55.806 | 52.435 | 17.619 | 70.818 | 189.486 | 52.702 | 18.566 | 75.683 | 238.988 |
| | | | Local | 41.268 | 20.601 | 58.996 | 27.027 | 61.259 | 31.348 | 59.187 | 27.09 | 67.386 | 82.774 |
| | | 0.5 | All | 59.966 | 69.622 | 134.113 | 40.508 | 153.303 | 197.561 | 152.764 | 194.281 | 159.805 | 254.390 |
| | | | Local | 53.404 | 37.459 | 151.241 | 73.12 | 155.068 | 74.906 | 154.86 | 74.706 | 162.092 | 105.748 |
| | | 1 | All | 69.85 | 71.463 | 235.957 | 69.586 | 257.054 | 211.767 | 256.694 | 210.392 | 264.881 | 270.685 |
| | | | Local | 65.513 | 49.217 | 264.334 | 129.118 | 270.066 | 131.648 | 269.859 | 131.429 | 280.553 | 158.214 |

Table C.5: (T2) Loading estimation errors of Trunc, iPE, RTFA and PreAve measured as in (C.2) for each mode scaled by 100, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | 1 | 1.949 | 0.302 | 2.197 | 0.335 | 2.194 | 0.335 | 4.528 | 0.977 |
| | | | 2 | 0.637 | 0.19 | 0.7 | 0.219 | 0.7 | 0.219 | 1.635 | 0.650 |
| | | | 3 | 0.631 | 0.165 | 0.69 | 0.156 | 0.69 | 0.156 | 1.579 | 0.494 |
| | | 0.5 | 1 | 2.073 | 0.305 | 3.122 | 0.452 | 3.117 | 0.451 | 6.471 | 1.558 |
| | | | 2 | 0.689 | 0.191 | 0.975 | 0.252 | 0.974 | 0.252 | 2.296 | 0.763 |
| | | | 3 | 0.681 | 0.176 | 0.955 | 0.243 | 0.953 | 0.243 | 2.296 | 0.755 |
| | | 1 | 1 | 2.197 | 0.295 | 3.972 | 0.541 | 3.968 | 0.54 | 8.227 | 1.691 |
| | | | 2 | 0.764 | 0.216 | 1.307 | 0.404 | 1.305 | 0.404 | 3.025 | 1.103 |
| | | | 3 | 0.737 | 0.183 | 1.241 | 0.302 | 1.239 | 0.302 | 2.964 | 0.897 |
| | $t_3$ | 0.1 | 1 | 2.037 | 0.284 | 2.595 | 0.396 | 2.583 | 0.394 | 5.632 | 1.295 |
| | | | 2 | 0.691 | 0.2 | 0.838 | 0.252 | 0.836 | 0.252 | 1.905 | 0.582 |
| | | | 3 | 0.681 | 0.218 | 0.832 | 0.257 | 0.829 | 0.257 | 2.023 | 0.671 |
| | | 0.5 | 1 | 2.31 | 0.333 | 4.212 | 0.528 | 4.201 | 0.527 | 9.186 | 2.544 |
| | | | 2 | 0.825 | 0.218 | 1.354 | 0.349 | 1.35 | 0.347 | 3.228 | 1.243 |
| | | | 3 | 0.842 | 0.27 | 1.325 | 0.416 | 1.322 | 0.418 | 3.153 | 0.933 |
| | | 1 | 1 | 2.568 | 0.433 | 5.682 | 0.761 | 5.672 | 0.758 | 13.284 | 4.824 |
| | | | 2 | 0.988 | 0.343 | 1.819 | 0.513 | 1.816 | 0.511 | 4.324 | 1.715 |
| | | | 3 | 1.013 | 0.357 | 1.835 | 0.492 | 1.833 | 0.491 | 4.428 | 1.492 |
| 200 | Gaussian | 0.1 | 1 | 1.438 | 0.188 | 1.634 | 0.207 | 1.631 | 0.206 | 3.321 | 0.696 |
| | | | 2 | 0.487 | 0.141 | 0.536 | 0.146 | 0.535 | 0.145 | 1.206 | 0.471 |
| | | | 3 | 0.463 | 0.125 | 0.517 | 0.14 | 0.515 | 0.14 | 1.169 | 0.313 |
| | | 0.5 | 1 | 1.541 | 0.188 | 2.347 | 0.267 | 2.343 | 0.267 | 4.738 | 0.825 |
| | | | 2 | 0.535 | 0.153 | 0.759 | 0.242 | 0.759 | 0.241 | 1.663 | 0.622 |
| | | | 3 | 0.501 | 0.128 | 0.716 | 0.171 | 0.714 | 0.17 | 1.658 | 0.480 |
| | | 1 | 1 | 1.628 | 0.191 | 2.966 | 0.33 | 2.963 | 0.329 | 6.047 | 1.004 |
| | | | 2 | 0.568 | 0.152 | 0.936 | 0.221 | 0.936 | 0.221 | 2.181 | 0.700 |
| | | | 3 | 0.548 | 0.141 | 0.945 | 0.237 | 0.943 | 0.237 | 2.194 | 0.571 |
| | $t_3$ | 0.1 | 1 | 1.445 | 0.188 | 1.912 | 0.245 | 1.9 | 0.242 | 5.136 | 8.030 |
| | | | 2 | 0.47 | 0.108 | 0.592 | 0.153 | 0.588 | 0.151 | 2.197 | 7.773 |
| | | | 3 | 0.478 | 0.125 | 0.598 | 0.141 | 0.592 | 0.139 | 2.173 | 7.713 |
| | | 0.5 | 1 | 1.673 | 0.299 | 3.277 | 0.42 | 3.267 | 0.42 | 7.996 | 8.033 |
| | | | 2 | 0.587 | 0.175 | 0.997 | 0.233 | 0.995 | 0.231 | 3.193 | 8.080 |
| | | | 3 | 0.582 | 0.184 | 1.015 | 0.262 | 1.014 | 0.26 | 3.333 | 8.098 |
| | | 1 | 1 | 1.86 | 0.37 | 4.446 | 0.584 | 4.438 | 0.584 | 11.622 | 11.228 |
| | | | 2 | 0.734 | 0.27 | 1.39 | 0.453 | 1.389 | 0.453 | 4.192 | 7.281 |
| | | | 3 | 0.727 | 0.265 | 1.436 | 0.432 | 1.433 | 0.432 | 4.457 | 8.761 |
| 500 | Gaussian | 0.1 | 1 | 0.895 | 0.124 | 1.019 | 0.136 | 1.017 | 0.137 | 1.977 | 0.334 |
| | | | 2 | 0.305 | 0.101 | 0.333 | 0.105 | 0.332 | 0.105 | 0.719 | 0.203 |
| | | | 3 | 0.313 | 0.103 | 0.342 | 0.107 | 0.342 | 0.107 | 0.724 | 0.212 |
| | | 0.5 | 1 | 0.964 | 0.125 | 1.471 | 0.18 | 1.468 | 0.179 | 2.835 | 0.432 |
| | | | 2 | 0.333 | 0.103 | 0.481 | 0.139 | 0.479 | 0.138 | 1.011 | 0.292 |
| | | | 3 | 0.343 | 0.104 | 0.495 | 0.141 | 0.494 | 0.141 | 1.049 | 0.278 |
| | | 1 | 1 | 1.037 | 0.129 | 1.904 | 0.221 | 1.902 | 0.22 | 3.667 | 0.620 |
| | | | 2 | 0.353 | 0.101 | 0.585 | 0.147 | 0.584 | 0.147 | 1.335 | 0.394 |
| | | | 3 | 0.369 | 0.114 | 0.628 | 0.202 | 0.627 | 0.202 | 1.396 | 0.387 |
| | $t_3$ | 0.1 | 1 | 0.973 | 0.174 | 1.345 | 0.194 | 1.336 | 0.193 | 2.727 | 0.654 |
| | | | 2 | 0.352 | 0.125 | 0.453 | 0.136 | 0.45 | 0.135 | 0.995 | 0.296 |
| | | | 3 | 0.336 | 0.101 | 0.435 | 0.129 | 0.432 | 0.129 | 0.935 | 0.347 |
| | | 0.5 | 1 | 1.202 | 0.32 | 2.475 | 0.399 | 2.466 | 0.396 | 4.987 | 1.673 |
| | | | 2 | 0.493 | 0.238 | 0.809 | 0.263 | 0.806 | 0.264 | 1.823 | 0.606 |
| | | | 3 | 0.457 | 0.195 | 0.752 | 0.197 | 0.75 | 0.198 | 1.792 | 1.293 |
| | | 1 | 1 | 1.386 | 0.378 | 3.466 | 0.525 | 3.459 | 0.524 | 7.296 | 4.967 |
| | | | 2 | 0.601 | 0.347 | 1.156 | 0.425 | 1.153 | 0.425 | 2.478 | 0.876 |
| | | | 3 | 0.574 | 0.376 | 1.052 | 0.327 | 1.05 | 0.328 | 2.487 | 1.848 |

Table C.6: (T2) Common component estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') scaled by 1000, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | noTrunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Range | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | All | 3.386 | 0.987 | 4.331 | 1.217 | 4.447 | 1.25 | 4.445 | 1.25 | 6.358 | 2.078 |
| | | | Local | 3.394 | 1.085 | 4.359 | 1.371 | 4.474 | 1.401 | 4.473 | 1.4 | 6.377 | 2.181 |
| | | 0.5 | All | 3.885 | 1.06 | 8.778 | 2.317 | 9.395 | 2.505 | 9.391 | 2.504 | 13.363 | 4.801 |
| | | | Local | 3.917 | 1.163 | 8.873 | 2.696 | 9.489 | 2.866 | 9.485 | 2.865 | 13.459 | 5.051 |
| | | 1 | All | 4.471 | 1.105 | 14.294 | 3.659 | 15.556 | 4.028 | 15.553 | 4.027 | 21.898 | 6.696 |
| | | | Local | 4.504 | 1.268 | 14.508 | 4.278 | 15.765 | 4.584 | 15.761 | 4.583 | 22.101 | 7.051 |
| | $t_3$ | 0.1 | All | 3.57 | 0.903 | 5.834 | 1.616 | 6.119 | 1.726 | 6.112 | 1.725 | 8.953 | 2.964 |
| | | | Local | 3.909 | 1.365 | 6.525 | 2.46 | 6.825 | 2.561 | 6.815 | 2.548 | 9.774 | 3.706 |
| | | 0.5 | All | 4.781 | 1.204 | 15.554 | 4.009 | 16.924 | 4.448 | 16.913 | 4.446 | 24.693 | 9.733 |
| | | | Local | 5.222 | 2.285 | 17.558 | 6.896 | 18.974 | 7.245 | 18.961 | 7.237 | 26.983 | 11.547 |
| | | 1 | All | 6.218 | 1.975 | 27.79 | 7.025 | 30.669 | 8 | 30.655 | 7.994 | 48.173 | 24.651 |
| | | | Local | 6.472 | 2.862 | 31.45 | 12.39 | 34.461 | 13.093 | 34.446 | 13.087 | 52.518 | 27.648 |
| 200 | Gaussian | 0.1 | All | 3.456 | 0.878 | 4.533 | 1.078 | 4.6 | 1.091 | 4.599 | 1.091 | 5.618 | 1.447 |
| | | | Local | 3.481 | 0.984 | 4.566 | 1.151 | 4.633 | 1.161 | 4.631 | 1.161 | 5.654 | 1.475 |
| | | 0.5 | All | 3.963 | 0.947 | 9.474 | 2.048 | 9.823 | 2.116 | 9.821 | 2.116 | 11.832 | 2.614 |
| | | | Local | 3.959 | 1.122 | 9.31 | 2.448 | 9.659 | 2.502 | 9.657 | 2.502 | 11.675 | 2.918 |
| | | 1 | All | 4.568 | 1.001 | 15.64 | 3.248 | 16.337 | 3.397 | 16.335 | 3.397 | 19.636 | 4.168 |
| | | | Local | 4.598 | 1.18 | 15.658 | 3.938 | 16.361 | 4.078 | 16.359 | 4.078 | 19.638 | 4.710 |
| | $t_3$ | 0.1 | All | 3.399 | 0.805 | 6.202 | 1.495 | 6.376 | 1.546 | 6.37 | 1.544 | 23.003 | 125.475 |
| | | | Local | 3.726 | 1.464 | 6.922 | 2.684 | 7.102 | 2.728 | 7.095 | 2.726 | 20.405 | 94.925 |
| | | 0.5 | All | 4.761 | 1.722 | 18.484 | 4.462 | 19.375 | 4.702 | 19.368 | 4.7 | 39.278 | 129.314 |
| | | | Local | 5.254 | 3.575 | 20.459 | 8.403 | 21.363 | 8.526 | 21.356 | 8.525 | 38.494 | 100.490 |
| | | 1 | All | 6.795 | 7.052 | 33.871 | 8.471 | 35.764 | 8.993 | 35.755 | 8.992 | 64.334 | 138.470 |
| | | | Local | 6.376 | 4.543 | 37.113 | 15.608 | 39.062 | 15.908 | 39.051 | 15.908 | 63.372 | 106.823 |
| 500 | Gaussian | 0.1 | All | 3.192 | 0.881 | 4.279 | 1.114 | 4.305 | 1.119 | 4.305 | 1.119 | 4.652 | 1.211 |
| | | | Local | 3.235 | 1.008 | 4.349 | 1.345 | 4.376 | 1.35 | 4.375 | 1.35 | 4.724 | 1.435 |
| | | 0.5 | All | 3.708 | 0.927 | 9.276 | 2.152 | 9.416 | 2.186 | 9.415 | 2.186 | 10.124 | 2.364 |
| | | | Local | 3.756 | 1.061 | 9.421 | 2.52 | 9.56 | 2.55 | 9.559 | 2.55 | 10.267 | 2.700 |
| | | 1 | All | 4.338 | 0.977 | 15.581 | 3.482 | 15.87 | 3.546 | 15.869 | 3.546 | 17.062 | 3.887 |
| | | | Local | 4.341 | 1.112 | 15.557 | 4.123 | 15.847 | 4.174 | 15.846 | 4.175 | 17.031 | 4.463 |
| | $t_3$ | 0.1 | All | 3.73 | 1.269 | 7.698 | 2.127 | 7.794 | 2.153 | 7.791 | 2.152 | 8.468 | 2.494 |
| | | | Local | 4.467 | 2.449 | 9.169 | 3.951 | 9.269 | 3.965 | 9.265 | 3.963 | 9.96 | 4.217 |
| | | 0.5 | All | 5.913 | 3.336 | 26.174 | 7.492 | 26.705 | 7.658 | 26.7 | 7.656 | 29.105 | 10.262 |
| | | | Local | 7.057 | 5.516 | 31.172 | 13.033 | 31.721 | 13.108 | 31.716 | 13.105 | 34.14 | 14.662 |
| | | 1 | All | 10.176 | 21.525 | 49.302 | 14.104 | 50.48 | 14.439 | 50.474 | 14.437 | 57.388 | 38.512 |
| | | | Local | 8.854 | 6.464 | 59.032 | 26.021 | 60.274 | 26.311 | 60.268 | 26.308 | 67.403 | 44.782 |

Table C.7: (T3) Loading estimation errors of Trunc, iPE, RTFA and PreAve measured as in (C.2) for each mode scaled by 100, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | 1 | 0.557 | 0.094 | 0.632 | 0.11 | 0.632 | 0.11 | 1.537 | 0.312 |
| | | | 2 | 0.689 | 0.093 | 0.78 | 0.103 | 0.78 | 0.104 | 1.86 | 0.291 |
| | | | 3 | 0.797 | 0.093 | 0.904 | 0.103 | 0.904 | 0.103 | 2.011 | 0.265 |
| | | 0.5 | 1 | 0.602 | 0.093 | 0.902 | 0.139 | 0.901 | 0.138 | 2.2 | 0.430 |
| | | | 2 | 0.748 | 0.093 | 1.122 | 0.147 | 1.122 | 0.147 | 2.655 | 0.380 |
| | | | 3 | 0.858 | 0.099 | 1.284 | 0.148 | 1.283 | 0.148 | 2.944 | 0.396 |
| | | 1 | 1 | 0.656 | 0.103 | 1.172 | 0.182 | 1.171 | 0.183 | 2.835 | 0.533 |
| | | | 2 | 0.808 | 0.097 | 1.442 | 0.173 | 1.441 | 0.173 | 3.449 | 0.488 |
| | | | 3 | 0.925 | 0.096 | 1.664 | 0.175 | 1.663 | 0.175 | 3.746 | 0.591 |
| | $t_3$ | 0.1 | 1 | 0.595 | 0.095 | 0.779 | 0.128 | 0.774 | 0.124 | 1.947 | 0.533 |
| | | | 2 | 0.741 | 0.112 | 0.988 | 0.148 | 0.981 | 0.145 | 2.877 | 5.488 |
| | | | 3 | 0.821 | 0.099 | 1.101 | 0.132 | 1.096 | 0.131 | 2.562 | 0.431 |
| | | 0.5 | 1 | 0.723 | 0.158 | 1.373 | 0.201 | 1.37 | 0.2 | 3.423 | 0.688 |
| | | | 2 | 0.873 | 0.181 | 1.659 | 0.248 | 1.656 | 0.247 | 4.603 | 5.479 |
| | | | 3 | 0.969 | 0.159 | 1.904 | 0.221 | 1.9 | 0.22 | 4.507 | 0.730 |
| | | 1 | 1 | 0.828 | 0.236 | 1.849 | 0.309 | 1.847 | 0.308 | 5.215 | 5.174 |
| | | | 2 | 0.999 | 0.247 | 2.25 | 0.3 | 2.247 | 0.299 | 6.17 | 5.254 |
| | | | 3 | 1.098 | 0.229 | 2.572 | 0.319 | 2.57 | 0.319 | 6.792 | 5.085 |
| 200 | Gaussian | 0.1 | 1 | 0.398 | 0.067 | 0.446 | 0.07 | 0.446 | 0.07 | 1.068 | 0.175 |
| | | | 2 | 0.493 | 0.06 | 0.558 | 0.061 | 0.558 | 0.062 | 1.284 | 0.195 |
| | | | 3 | 0.578 | 0.071 | 0.657 | 0.078 | 0.656 | 0.078 | 1.481 | 0.237 |
| | | 0.5 | 1 | 0.44 | 0.072 | 0.659 | 0.096 | 0.659 | 0.096 | 1.544 | 0.249 |
| | | | 2 | 0.535 | 0.07 | 0.806 | 0.105 | 0.806 | 0.106 | 1.868 | 0.263 |
| | | | 3 | 0.633 | 0.068 | 0.957 | 0.104 | 0.957 | 0.104 | 2.164 | 0.299 |
| | | 1 | 1 | 0.48 | 0.074 | 0.853 | 0.126 | 0.852 | 0.126 | 2.023 | 0.335 |
| | | | 2 | 0.587 | 0.068 | 1.062 | 0.131 | 1.062 | 0.13 | 2.375 | 0.347 |
| | | | 3 | 0.683 | 0.069 | 1.223 | 0.124 | 1.222 | 0.124 | 2.77 | 0.404 |
| | $t_3$ | 0.1 | 1 | 0.437 | 0.081 | 0.605 | 0.092 | 0.603 | 0.092 | 1.447 | 0.405 |
| | | | 2 | 0.531 | 0.099 | 0.739 | 0.112 | 0.736 | 0.112 | 1.719 | 0.295 |
| | | | 3 | 0.606 | 0.098 | 0.854 | 0.13 | 0.852 | 0.129 | 1.977 | 0.337 |
| | | 0.5 | 1 | 0.563 | 0.226 | 1.104 | 0.238 | 1.102 | 0.238 | 2.682 | 0.819 |
| | | | 2 | 0.672 | 0.227 | 1.336 | 0.234 | 1.334 | 0.234 | 3.192 | 0.629 |
| | | | 3 | 0.759 | 0.209 | 1.554 | 0.229 | 1.552 | 0.227 | 3.63 | 0.661 |
| | | 1 | 1 | 0.673 | 0.297 | 1.525 | 0.347 | 1.523 | 0.346 | 3.842 | 1.160 |
| | | | 2 | 0.791 | 0.275 | 1.859 | 0.258 | 1.858 | 0.257 | 4.585 | 1.024 |
| | | | 3 | 0.892 | 0.272 | 2.133 | 0.291 | 2.131 | 0.291 | 5.083 | 1.042 |
| 500 | Gaussian | 0.1 | 1 | 0.253 | 0.033 | 0.29 | 0.04 | 0.29 | 0.04 | 0.666 | 0.101 |
| | | | 2 | 0.315 | 0.04 | 0.358 | 0.045 | 0.358 | 0.044 | 0.816 | 0.116 |
| | | | 3 | 0.371 | 0.043 | 0.42 | 0.048 | 0.419 | 0.048 | 0.944 | 0.125 |
| | | 0.5 | 1 | 0.275 | 0.033 | 0.416 | 0.053 | 0.416 | 0.052 | 0.976 | 0.143 |
| | | | 2 | 0.345 | 0.041 | 0.524 | 0.06 | 0.524 | 0.06 | 1.191 | 0.143 |
| | | | 3 | 0.407 | 0.045 | 0.611 | 0.064 | 0.611 | 0.064 | 1.359 | 0.174 |
| | | 1 | 1 | 0.3 | 0.038 | 0.545 | 0.068 | 0.545 | 0.068 | 1.256 | 0.182 |
| | | | 2 | 0.379 | 0.049 | 0.683 | 0.09 | 0.682 | 0.09 | 1.545 | 0.212 |
| | | | 3 | 0.441 | 0.045 | 0.789 | 0.083 | 0.788 | 0.083 | 1.766 | 0.203 |
| | $t_3$ | 0.1 | 1 | 0.285 | 0.056 | 0.427 | 0.059 | 0.425 | 0.059 | 1.01 | 0.182 |
| | | | 2 | 0.337 | 0.056 | 0.513 | 0.059 | 0.511 | 0.059 | 1.163 | 0.167 |
| | | | 3 | 0.396 | 0.065 | 0.604 | 0.064 | 0.601 | 0.064 | 1.327 | 0.195 |
| | | 0.5 | 1 | 0.381 | 0.109 | 0.819 | 0.123 | 0.818 | 0.122 | 1.968 | 0.310 |
| | | | 2 | 0.456 | 0.14 | 1.012 | 0.129 | 1.011 | 0.128 | 2.334 | 0.400 |
| | | | 3 | 0.525 | 0.165 | 1.162 | 0.176 | 1.16 | 0.176 | 2.561 | 0.382 |
| | | 1 | 1 | 0.468 | 0.169 | 1.158 | 0.193 | 1.157 | 0.193 | 2.776 | 0.535 |
| | | | 2 | 0.552 | 0.213 | 1.419 | 0.181 | 1.418 | 0.181 | 3.288 | 0.523 |
| | | | 3 | 0.63 | 0.234 | 1.623 | 0.246 | 1.622 | 0.246 | 3.687 | 0.618 |

Table C.8: (T3) Common component estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') scaled by 1000, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | noTrunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Range | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | All | 1.351 | 0.254 | 1.766 | 0.32 | 1.803 | 0.326 | 1.803 | 0.326 | 2.533 | 0.476 |
| | | | Local | 1.354 | 0.3 | 1.772 | 0.395 | 1.809 | 0.4 | 1.808 | 0.4 | 2.541 | 0.537 |
| | | 0.5 | All | 1.607 | 0.277 | 3.741 | 0.627 | 3.929 | 0.66 | 3.929 | 0.66 | 5.465 | 0.967 |
| | | | Local | 1.621 | 0.32 | 3.749 | 0.763 | 3.937 | 0.788 | 3.937 | 0.788 | 5.468 | 1.078 |
| | | 1 | All | 1.938 | 0.297 | 6.204 | 0.968 | 6.594 | 1.036 | 6.593 | 1.036 | 9.135 | 1.563 |
| | | | Local | 1.948 | 0.352 | 6.188 | 1.246 | 6.578 | 1.298 | 6.578 | 1.298 | 9.126 | 1.810 |
| | $t_3$ | 0.1 | All | 1.408 | 0.27 | 2.628 | 0.559 | 2.738 | 0.589 | 2.734 | 0.585 | 6.104 | 22.733 |
| | | | Local | 1.549 | 0.59 | 2.865 | 0.946 | 2.975 | 0.96 | 2.972 | 0.96 | 6.371 | 22.330 |
| | | 0.5 | All | 2.022 | 0.577 | 7.957 | 1.554 | 8.491 | 1.683 | 8.488 | 1.682 | 14.171 | 22.377 |
| | | | Local | 2.24 | 1.306 | 8.715 | 3.145 | 9.262 | 3.196 | 9.259 | 3.195 | 15.064 | 21.899 |
| | | 1 | All | 2.741 | 1.007 | 14.682 | 2.986 | 15.767 | 3.257 | 15.764 | 3.256 | 29.664 | 70.941 |
| | | | Local | 3.025 | 1.996 | 16.476 | 5.812 | 17.614 | 5.954 | 17.611 | 5.954 | 28.85 | 41.863 |
| 200 | Gaussian | 0.1 | All | 1.342 | 0.273 | 1.788 | 0.342 | 1.807 | 0.345 | 1.807 | 0.344 | 2.184 | 0.434 |
| | | | Local | 1.357 | 0.343 | 1.818 | 0.446 | 1.837 | 0.449 | 1.837 | 0.449 | 2.217 | 0.528 |
| | | 0.5 | All | 1.595 | 0.284 | 3.868 | 0.656 | 3.971 | 0.674 | 3.971 | 0.674 | 4.765 | 0.841 |
| | | | Local | 1.602 | 0.363 | 3.898 | 0.848 | 4.001 | 0.862 | 4 | 0.862 | 4.8 | 1.006 |
| | | 1 | All | 1.937 | 0.307 | 6.538 | 1.04 | 6.749 | 1.075 | 6.749 | 1.075 | 8.05 | 1.343 |
| | | | Local | 1.947 | 0.394 | 6.617 | 1.354 | 6.828 | 1.378 | 6.827 | 1.378 | 8.136 | 1.602 |
| | $t_3$ | 0.1 | All | 1.489 | 0.36 | 3.103 | 0.6 | 3.174 | 0.617 | 3.173 | 0.617 | 3.81 | 0.741 |
| | | | Local | 1.733 | 0.996 | 3.534 | 1.462 | 3.608 | 1.462 | 3.607 | 1.462 | 4.275 | 1.657 |
| | | 0.5 | All | 2.398 | 1.241 | 10.437 | 2.149 | 10.795 | 2.216 | 10.794 | 2.215 | 12.995 | 2.735 |
| | | | Local | 2.942 | 3.15 | 11.998 | 5.899 | 12.374 | 5.945 | 12.372 | 5.944 | 14.677 | 6.936 |
| | | 1 | All | 3.388 | 2.15 | 19.624 | 3.984 | 20.379 | 4.123 | 20.378 | 4.122 | 24.875 | 5.427 |
| | | | Local | 4.144 | 4.844 | 22.629 | 10.679 | 23.418 | 10.742 | 23.416 | 10.741 | 28.171 | 12.722 |
| 500 | Gaussian | 0.1 | All | 1.299 | 0.219 | 1.766 | 0.278 | 1.774 | 0.279 | 1.774 | 0.279 | 1.923 | 0.301 |
| | | | Local | 1.322 | 0.294 | 1.793 | 0.401 | 1.801 | 0.402 | 1.801 | 0.402 | 1.95 | 0.424 |
| | | 0.5 | All | 1.571 | 0.231 | 3.976 | 0.542 | 4.019 | 0.548 | 4.019 | 0.548 | 4.333 | 0.586 |
| | | | Local | 1.591 | 0.306 | 4.022 | 0.73 | 4.065 | 0.734 | 4.064 | 0.734 | 4.38 | 0.766 |
| | | 1 | All | 1.926 | 0.255 | 6.774 | 0.88 | 6.862 | 0.893 | 6.862 | 0.893 | 7.388 | 0.962 |
| | | | Local | 1.953 | 0.352 | 6.841 | 1.325 | 6.928 | 1.335 | 6.928 | 1.335 | 7.456 | 1.400 |
| | $t_3$ | 0.1 | All | 1.541 | 0.443 | 3.894 | 0.696 | 3.935 | 0.703 | 3.935 | 0.702 | 4.232 | 0.751 |
| | | | Local | 1.74 | 0.844 | 4.378 | 1.624 | 4.42 | 1.628 | 4.419 | 1.628 | 4.722 | 1.659 |
| | | 0.5 | All | 2.776 | 1.715 | 14.781 | 2.767 | 15.002 | 2.81 | 15.001 | 2.81 | 16.144 | 2.983 |
| | | | Local | 3.131 | 2.438 | 16.71 | 6.14 | 16.938 | 6.181 | 16.937 | 6.181 | 18.099 | 6.319 |
| | | 1 | All | 4.107 | 3.145 | 28.423 | 5.344 | 28.894 | 5.426 | 28.893 | 5.426 | 31.221 | 5.884 |
| | | | Local | 4.518 | 4.241 | 31.646 | 11.523 | 32.147 | 11.617 | 32.145 | 11.617 | 34.526 | 11.974 |

**Outliers in the factors.** See Figures C.9–C.14 and Tables C.9–C.14 for the results from loading and common component estimation obtained under (T1)–(T3) with outliers in the factors under (O2). In this scenario, we cannot recover $\mathcal{F}_t^\circ$ prior to the contamination by outliers since all cross-sections of the observed $\mathcal{X}_t$ are contaminated by the outliers. Nonetheless, we observe that the loadings and $\boldsymbol{\chi}_t$ (post-contamination) are well-estimated by Trunc. In line with the theory (Theorem 4 (i)), the estimation error tends to increase with the increase in the proportion of outliers.

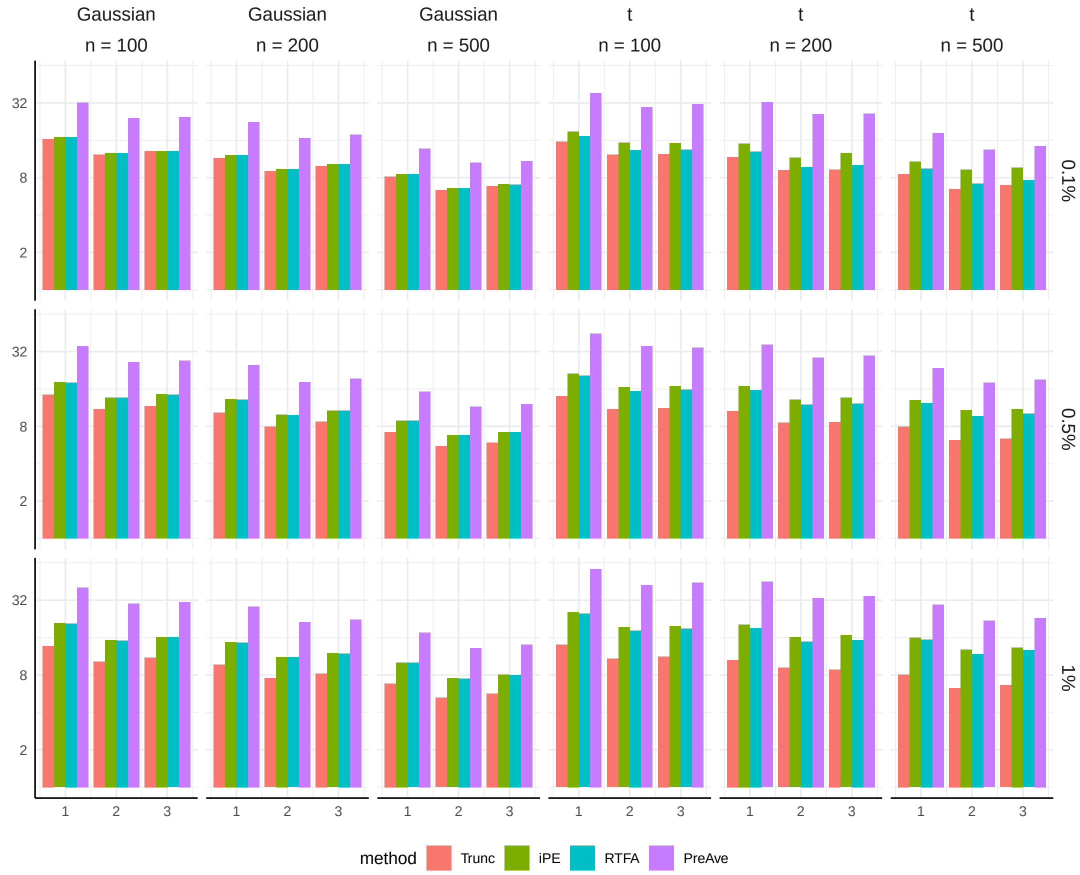


Figure C.9: (T1) Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

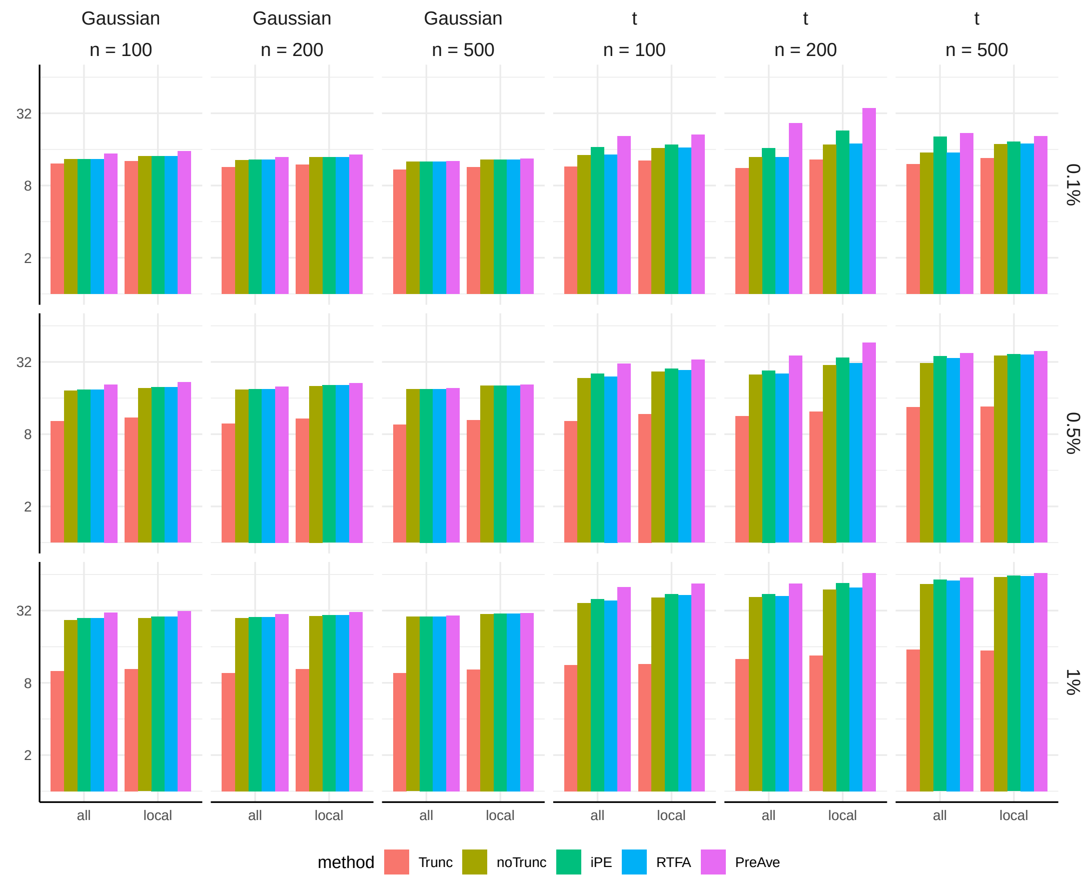


Figure C.10: (T1) Common component estimation errors measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') for Trunc, noTrunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

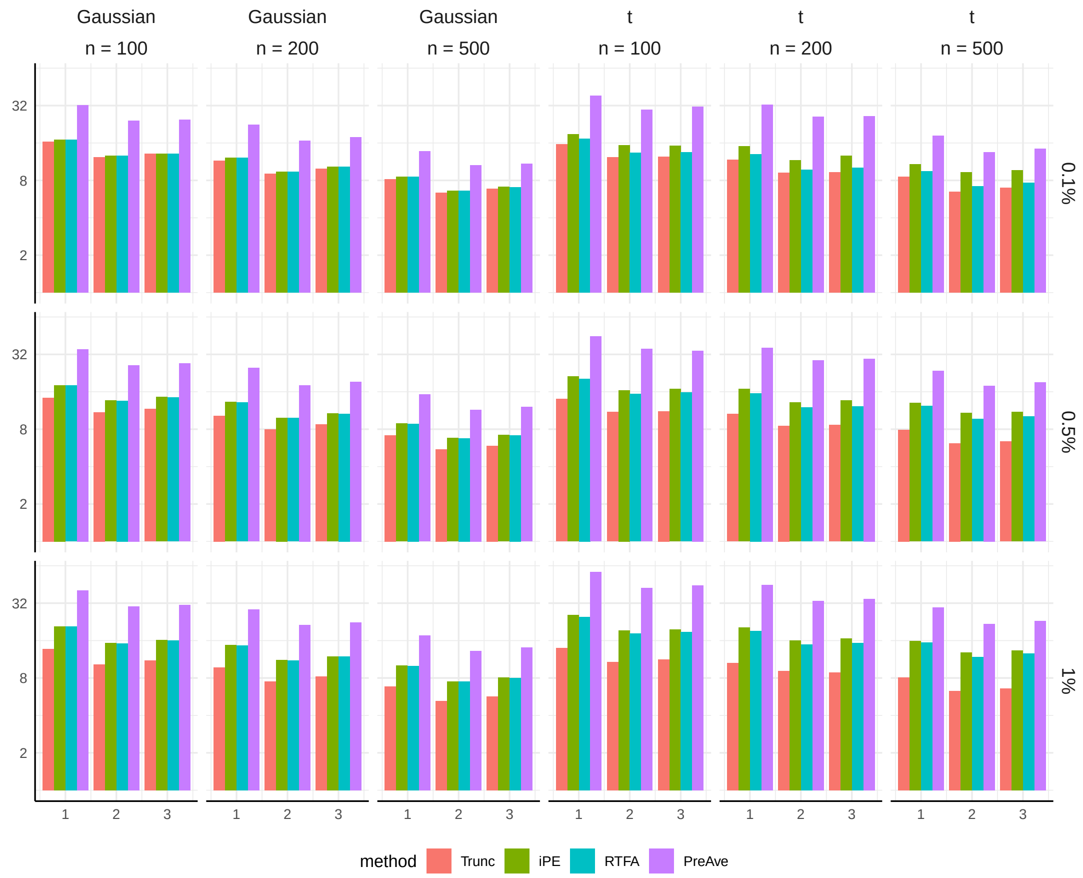


Figure C.11: (T2) Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

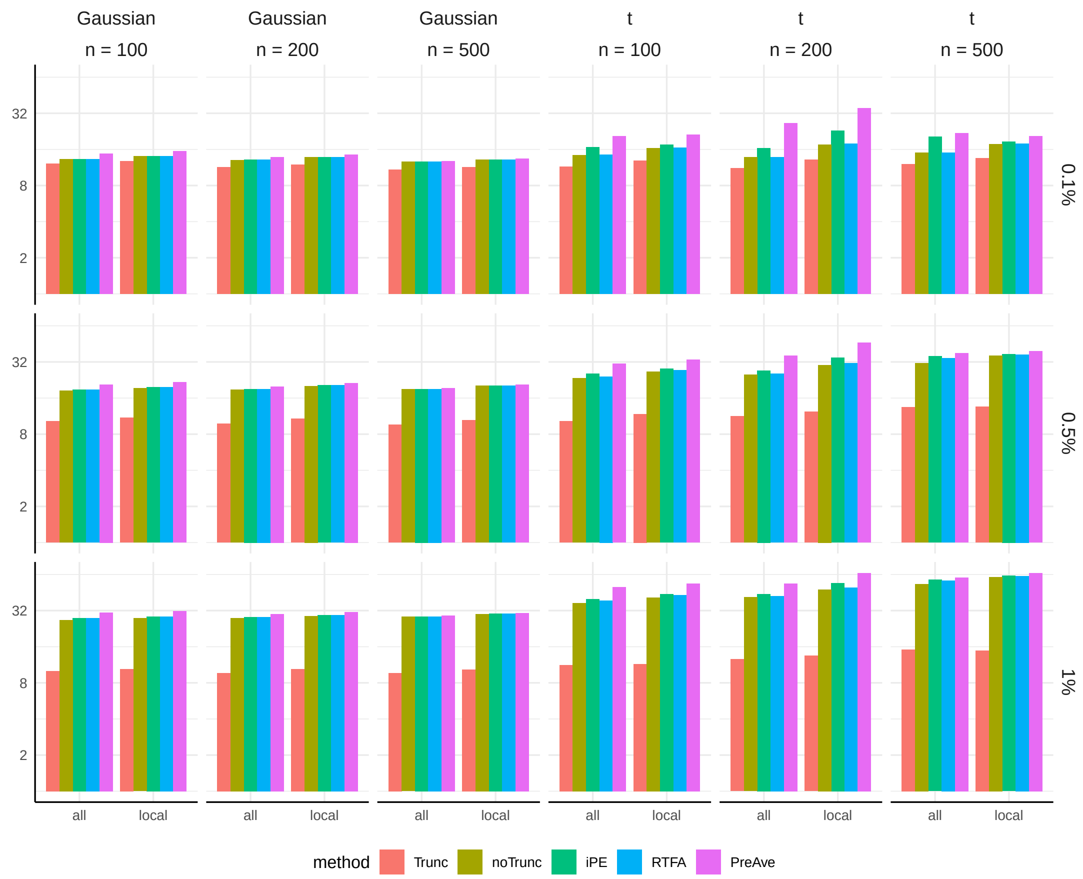


Figure C.12: (T2) Common component estimation errors measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') for Trunc, noTrunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

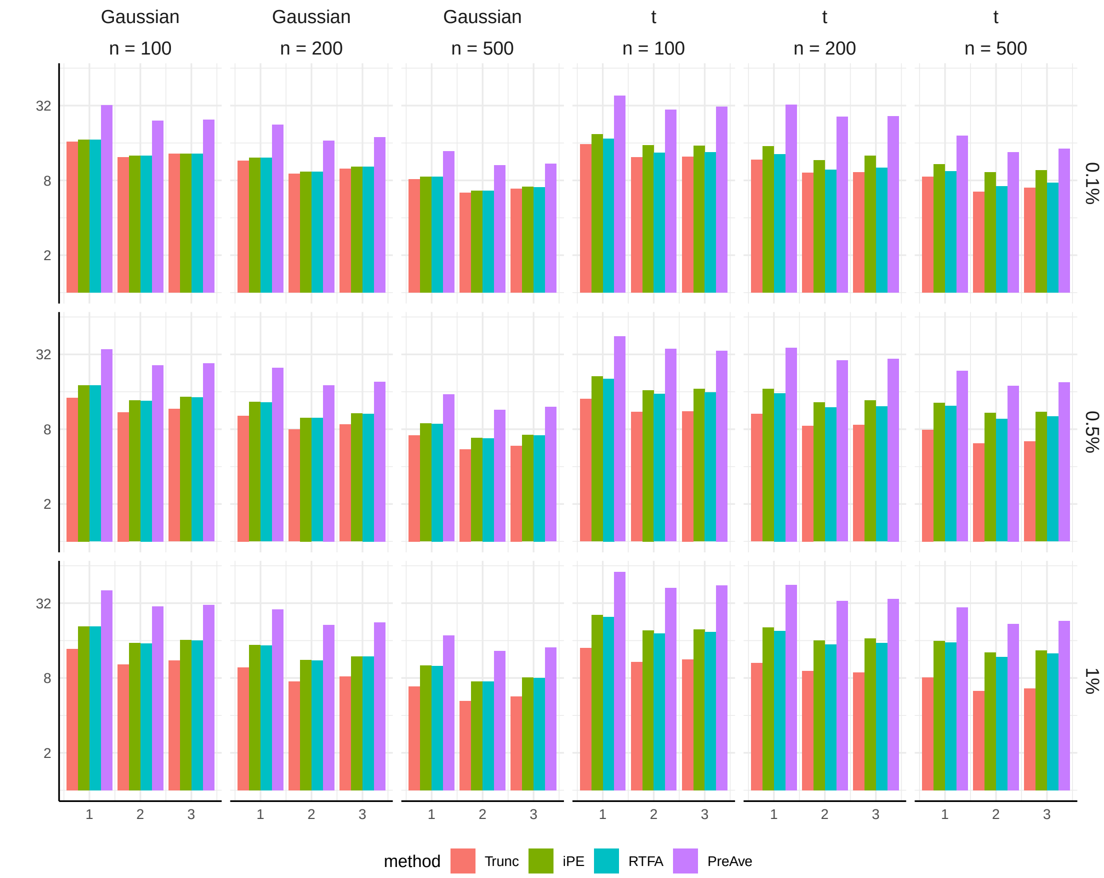


Figure C.13: (T3) Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

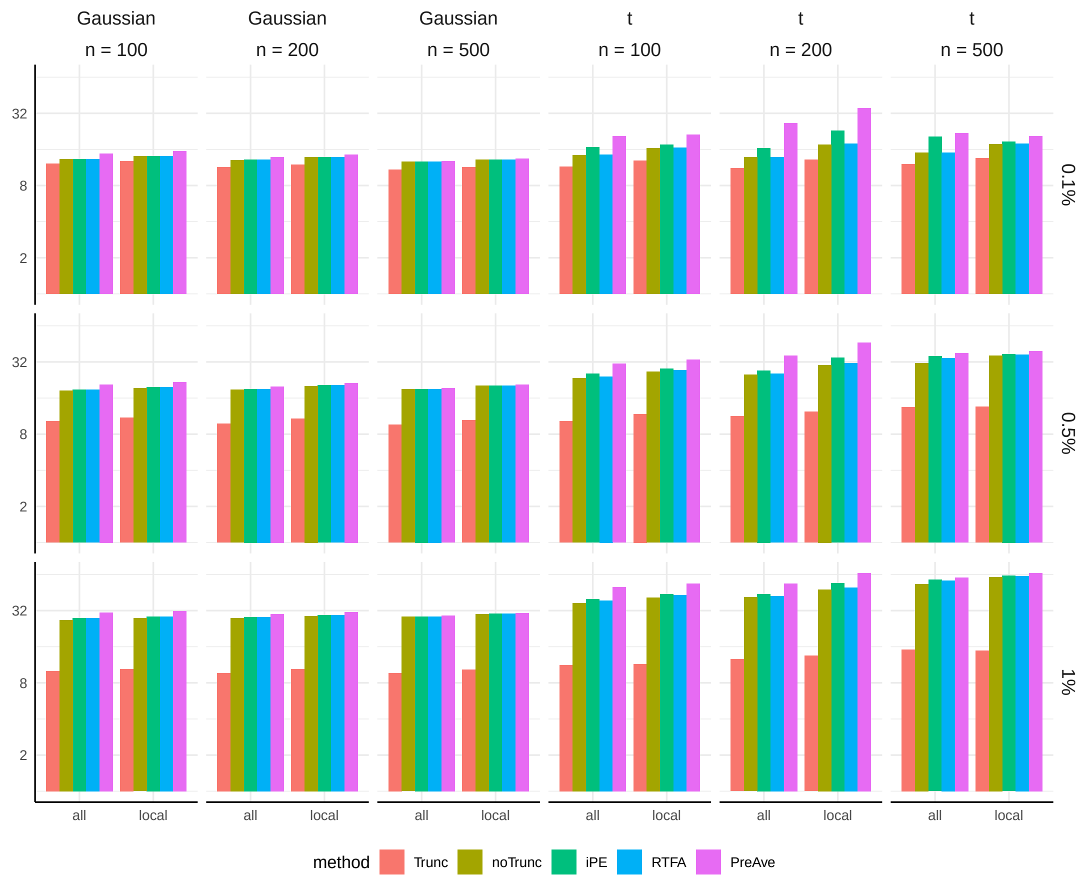


Figure C.14: (T3) Common component estimation errors measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') for Trunc, noTrunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers ($\varrho \in \{0.1, 0.5, 1\}$, top to bottom), averaged over 100 realisations per setting. In each plot, the $y$-axis is in the log-scale and all errors have been scaled for the ease of presentation.

Table C.9: (T1) Loading estimation errors of Trunc, iPE, RTFA and PreAve measured as in (C.2) for each mode scaled by 100, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | 1 | 2.415 | 1.876 | 2.278 | 1.034 | 2.279 | 1.037 | 3.875 | 1.545 |
| | | | 2 | 2.326 | 1.32 | 2.238 | 1.08 | 2.238 | 1.079 | 3.792 | 1.253 |
| | | | 3 | 2.552 | 3.733 | 2.198 | 0.966 | 2.197 | 0.964 | 3.764 | 1.434 |
| | | 0.5 | 1 | 1.507 | 0.651 | 1.471 | 0.543 | 1.471 | 0.545 | 2.692 | 0.977 |
| | | | 2 | 1.555 | 0.775 | 1.523 | 0.72 | 1.523 | 0.719 | 2.605 | 0.933 |
| | | | 3 | 1.493 | 0.642 | 1.471 | 0.587 | 1.471 | 0.587 | 2.584 | 0.845 |
| | | 1 | 1 | 1.145 | 0.431 | 1.14 | 0.43 | 1.14 | 0.433 | 2.073 | 0.741 |
| | | | 2 | 1.121 | 0.414 | 1.116 | 0.412 | 1.116 | 0.412 | 1.889 | 0.572 |
| | | | 3 | 1.121 | 0.502 | 1.119 | 0.501 | 1.119 | 0.501 | 2.05 | 0.750 |
| | $t_3$ | 0.1 | 1 | 2.075 | 0.797 | 2.416 | 3.66 | 2.01 | 0.782 | 4.159 | 4.215 |
| | | | 2 | 2.216 | 0.966 | 2.729 | 5.495 | 2.127 | 0.924 | 4.39 | 5.347 |
| | | | 3 | 2.209 | 0.847 | 2.606 | 4.508 | 2.101 | 0.727 | 4.541 | 5.134 |
| | | 0.5 | 1 | 1.195 | 0.448 | 1.194 | 0.425 | 1.166 | 0.4 | 2.442 | 1.025 |
| | | | 2 | 1.295 | 0.472 | 1.312 | 0.549 | 1.263 | 0.462 | 2.586 | 1.586 |
| | | | 3 | 1.234 | 0.398 | 1.248 | 0.427 | 1.203 | 0.358 | 2.578 | 1.561 |
| | | 1 | 1 | 0.851 | 0.23 | 0.862 | 0.25 | 0.842 | 0.229 | 1.753 | 0.673 |
| | | | 2 | 0.884 | 0.284 | 0.914 | 0.383 | 0.88 | 0.289 | 1.786 | 0.747 |
| | | | 3 | 0.885 | 0.298 | 0.896 | 0.309 | 0.874 | 0.29 | 1.802 | 0.687 |
| 200 | Gaussian | 0.1 | 1 | 1.585 | 0.694 | 1.564 | 0.678 | 1.565 | 0.679 | 2.509 | 0.891 |
| | | | 2 | 1.656 | 0.695 | 1.645 | 0.688 | 1.645 | 0.689 | 2.498 | 0.808 |
| | | | 3 | 1.831 | 1.385 | 1.811 | 1.345 | 1.811 | 1.347 | 2.613 | 1.484 |
| | | 0.5 | 1 | 1.009 | 0.397 | 1.005 | 0.396 | 1.005 | 0.397 | 1.784 | 0.525 |
| | | | 2 | 1.072 | 0.387 | 1.068 | 0.384 | 1.069 | 0.385 | 1.71 | 0.538 |
| | | | 3 | 1.156 | 0.76 | 1.15 | 0.756 | 1.151 | 0.757 | 1.809 | 0.658 |
| | | 1 | 1 | 0.78 | 0.257 | 0.778 | 0.258 | 0.778 | 0.257 | 1.343 | 0.404 |
| | | | 2 | 0.778 | 0.258 | 0.777 | 0.257 | 0.776 | 0.257 | 1.366 | 0.405 |
| | | | 3 | 0.852 | 0.454 | 0.849 | 0.451 | 0.849 | 0.452 | 1.384 | 0.522 |
| | $t_3$ | 0.1 | 1 | 1.576 | 0.758 | 2.038 | 3.59 | 1.534 | 0.72 | 3.826 | 9.825 |
| | | | 2 | 1.593 | 0.718 | 1.918 | 2.637 | 1.533 | 0.644 | 4.183 | 10.317 |
| | | | 3 | 1.545 | 0.581 | 2.19 | 5.65 | 1.505 | 0.565 | 3.518 | 5.420 |
| | | 0.5 | 1 | 0.897 | 0.419 | 0.931 | 0.506 | 0.881 | 0.414 | 1.555 | 0.630 |
| | | | 2 | 0.896 | 0.347 | 0.928 | 0.517 | 0.871 | 0.334 | 1.616 | 0.719 |
| | | | 3 | 0.886 | 0.277 | 0.928 | 0.494 | 0.869 | 0.272 | 1.566 | 0.615 |
| | | 1 | 1 | 0.637 | 0.192 | 0.649 | 0.237 | 0.625 | 0.186 | 1.136 | 0.390 |
| | | | 2 | 0.648 | 0.229 | 0.667 | 0.3 | 0.638 | 0.225 | 1.216 | 0.560 |
| | | | 3 | 0.65 | 0.195 | 0.662 | 0.235 | 0.64 | 0.194 | 1.135 | 0.358 |
| 500 | Gaussian | 0.1 | 1 | 1.251 | 0.655 | 1.245 | 0.647 | 1.245 | 0.648 | 1.558 | 0.568 |
| | | | 2 | 1.188 | 0.546 | 1.182 | 0.536 | 1.181 | 0.536 | 1.579 | 0.512 |
| | | | 3 | 1.285 | 0.743 | 1.279 | 0.74 | 1.279 | 0.741 | 1.576 | 0.467 |
| | | 0.5 | 1 | 0.771 | 0.325 | 0.767 | 0.323 | 0.767 | 0.323 | 1.073 | 0.330 |
| | | | 2 | 0.753 | 0.277 | 0.751 | 0.276 | 0.75 | 0.276 | 1.099 | 0.347 |
| | | | 3 | 0.767 | 0.373 | 0.766 | 0.373 | 0.766 | 0.373 | 1.066 | 0.295 |
| | | 1 | 1 | 0.544 | 0.224 | 0.543 | 0.223 | 0.542 | 0.223 | 0.852 | 0.274 |
| | | | 2 | 0.544 | 0.202 | 0.543 | 0.202 | 0.543 | 0.202 | 0.848 | 0.268 |
| | | | 3 | 0.581 | 0.256 | 0.579 | 0.256 | 0.579 | 0.256 | 0.838 | 0.242 |
| | $t_3$ | 0.1 | 1 | 1.22 | 0.709 | 1.58 | 3.408 | 1.191 | 0.669 | 2.197 | 6.414 |
| | | | 2 | 1.13 | 0.625 | 1.713 | 5.478 | 1.104 | 0.605 | 2.012 | 5.513 |
| | | | 3 | 1.23 | 0.875 | 1.768 | 5.127 | 1.195 | 0.835 | 2.15 | 5.979 |
| | | 0.5 | 1 | 0.667 | 0.299 | 0.944 | 2.692 | 0.655 | 0.29 | 1.415 | 4.459 |
| | | | 2 | 0.607 | 0.278 | 1.168 | 5.422 | 0.599 | 0.277 | 1.456 | 5.517 |
| | | | 3 | 0.651 | 0.376 | 1.049 | 3.83 | 0.64 | 0.355 | 1.48 | 5.144 |
| | | 1 | 1 | 0.45 | 0.174 | 0.655 | 2.008 | 0.442 | 0.17 | 0.732 | 0.233 |
| | | | 2 | 0.425 | 0.149 | 0.946 | 5.059 | 0.422 | 0.143 | 1.057 | 3.531 |
| | | | 3 | 0.441 | 0.229 | 0.707 | 2.561 | 0.436 | 0.22 | 0.73 | 0.305 |

Table C.10: (T1) Common component estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n-10+1, \ldots, n\}$ ('local') scaled by 1000, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | noTrunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Range | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | All | 30.938 | 40.311 | 27.567 | 10.331 | 27.09 | 8.464 | 27.089 | 8.464 | 29.024 | 9.080 |
| | | | Local | 33.127 | 35.111 | 30.493 | 13.661 | 29.976 | 11.932 | 29.976 | 11.932 | 31.914 | 12.370 |
| | | 0.5 | All | 14.614 | 5.135 | 14.48 | 4.689 | 14.445 | 4.633 | 14.445 | 4.633 | 15.398 | 4.940 |
| | | | Local | 17.564 | 8.874 | 17.566 | 8.878 | 17.535 | 8.829 | 17.534 | 8.828 | 18.486 | 8.983 |
| | | 1 | All | 9.032 | 2.763 | 9.027 | 2.763 | 9.024 | 2.763 | 9.024 | 2.763 | 9.6 | 2.940 |
| | | | Local | 10.613 | 5.993 | 10.612 | 5.996 | 10.61 | 5.995 | 10.61 | 5.995 | 11.2 | 6.085 |
| | $t_3$ | 0.1 | All | 26.006 | 12.283 | 26.411 | 12.607 | 31.938 | 57.573 | 26.277 | 12.562 | 34.995 | 65.368 |
| | | | Local | 30.569 | 17.588 | 31.794 | 20.747 | 33.713 | 26.996 | 31.722 | 20.87 | 36.316 | 29.472 |
| | | 0.5 | All | 11.163 | 4.834 | 11.368 | 4.935 | 11.401 | 4.953 | 11.342 | 4.904 | 12.705 | 6.752 |
| | | | Local | 16.349 | 11.237 | 16.642 | 11.726 | 16.658 | 11.737 | 16.616 | 11.731 | 17.893 | 11.960 |
| | | 1 | All | 6.49 | 2.787 | 6.589 | 2.847 | 6.603 | 2.856 | 6.583 | 2.845 | 7.15 | 3.130 |
| | | | Local | 7.743 | 5.667 | 7.827 | 5.783 | 7.835 | 5.789 | 7.821 | 5.781 | 8.407 | 5.899 |
| 200 | Gaussian | 0.1 | All | 26.147 | 10.189 | 26.122 | 10.135 | 26.1 | 10.098 | 26.101 | 10.099 | 26.777 | 10.272 |
| | | | Local | 28.654 | 14.543 | 28.655 | 14.55 | 28.633 | 14.509 | 28.633 | 14.511 | 29.281 | 14.556 |
| | | 0.5 | All | 14.03 | 5.16 | 14.022 | 5.156 | 14.02 | 5.153 | 14.02 | 5.153 | 14.396 | 5.150 |
| | | | Local | 18.545 | 11.559 | 18.544 | 11.56 | 18.542 | 11.559 | 18.542 | 11.559 | 18.917 | 11.585 |
| | | 1 | All | 8.935 | 3.246 | 8.931 | 3.246 | 8.93 | 3.245 | 8.93 | 3.245 | 9.179 | 3.284 |
| | | | Local | 12.311 | 7.982 | 12.309 | 7.984 | 12.308 | 7.983 | 12.308 | 7.983 | 12.553 | 8.005 |
| | $t_3$ | 0.1 | All | 22.249 | 6.565 | 23.004 | 7.664 | 28.682 | 52.623 | 22.982 | 7.797 | 42.115 | 122.632 |
| | | | Local | 29.79 | 16.019 | 31.991 | 23.092 | 38.13 | 57.329 | 32.165 | 24.722 | 52.687 | 127.384 |
| | | 0.5 | All | 10.089 | 3.136 | 10.339 | 3.539 | 10.408 | 3.932 | 10.329 | 3.54 | 10.689 | 3.681 |
| | | | Local | 15.83 | 12.669 | 16.684 | 15.145 | 17.041 | 17.383 | 16.682 | 15.194 | 17.04 | 15.154 |
| | | 1 | All | 6.055 | 1.787 | 6.188 | 2.044 | 6.206 | 2.144 | 6.184 | 2.046 | 6.401 | 2.259 |
| | | | Local | 9.413 | 8.219 | 10.249 | 12.98 | 10.414 | 14.311 | 10.255 | 13.05 | 10.645 | 14.730 |
| 500 | Gaussian | 0.1 | All | 24.782 | 7.499 | 24.775 | 7.495 | 24.77 | 7.49 | 24.77 | 7.49 | 24.824 | 7.319 |
| | | | Local | 27.414 | 11.115 | 27.414 | 11.114 | 27.41 | 11.113 | 27.41 | 11.113 | 27.451 | 10.997 |
| | | 0.5 | All | 13.626 | 4.076 | 13.621 | 4.075 | 13.621 | 4.074 | 13.62 | 4.074 | 13.704 | 4.042 |
| | | | Local | 17.421 | 9.794 | 17.42 | 9.795 | 17.419 | 9.795 | 17.419 | 9.795 | 17.499 | 9.777 |
| | | 1 | All | 8.63 | 2.561 | 8.626 | 2.561 | 8.626 | 2.561 | 8.626 | 2.561 | 8.7 | 2.564 |
| | | | Local | 11.029 | 6.947 | 11.028 | 6.948 | 11.027 | 6.949 | 11.027 | 6.949 | 11.098 | 6.936 |
| | $t_3$ | 0.1 | All | 21.868 | 7.997 | 22.755 | 9.187 | 36.51 | 141.46 | 22.751 | 9.252 | 40.021 | 176.553 |
| | | | Local | 29.237 | 16.434 | 29.871 | 16.804 | 32.561 | 29.594 | 29.848 | 16.788 | 36.738 | 68.349 |
| | | 0.5 | All | 10.172 | 3.431 | 10.512 | 3.923 | 17.669 | 73.514 | 10.511 | 3.933 | 20.002 | 96.155 |
| | | | Local | 15.185 | 13.995 | 15.343 | 14.071 | 17.284 | 23.15 | 15.338 | 14.067 | 19.292 | 40.375 |
| | | 1 | All | 6.074 | 2.029 | 6.258 | 2.294 | 10.023 | 38.78 | 6.257 | 2.292 | 7.256 | 10.680 |
| | | | Local | 9.216 | 7.626 | 9.29 | 7.645 | 10.984 | 18.065 | 9.289 | 7.643 | 9.864 | 8.892 |

Table C.11: (T2) Loading estimation errors of Trunc, iPE, RTFA and PreAve measured as in (C.2) for each mode scaled by 100, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | 1 | 1.766 | 0.275 | 1.765 | 0.274 | 1.765 | 0.274 | 3.565 | 0.777 |
| | | | 2 | 0.562 | 0.164 | 0.558 | 0.163 | 0.558 | 0.163 | 1.331 | 0.437 |
| | | | 3 | 0.56 | 0.155 | 0.556 | 0.154 | 0.556 | 0.154 | 1.307 | 0.454 |
| | | 0.5 | 1 | 1.271 | 0.208 | 1.269 | 0.208 | 1.269 | 0.208 | 2.632 | 0.637 |
| | | | 2 | 0.406 | 0.111 | 0.402 | 0.111 | 0.402 | 0.111 | 0.939 | 0.344 |
| | | | 3 | 0.391 | 0.101 | 0.39 | 0.102 | 0.39 | 0.102 | 0.922 | 0.348 |
| | | 1 | 1 | 0.994 | 0.161 | 0.993 | 0.161 | 0.993 | 0.161 | 2.032 | 0.515 |
| | | | 2 | 0.32 | 0.086 | 0.319 | 0.085 | 0.319 | 0.085 | 0.759 | 0.351 |
| | | | 3 | 0.302 | 0.078 | 0.3 | 0.077 | 0.3 | 0.077 | 0.682 | 0.207 |
| | $t_3$ | 0.1 | 1 | 1.745 | 0.285 | 1.776 | 0.3 | 1.768 | 0.297 | 3.897 | 1.109 |
| | | | 2 | 0.562 | 0.176 | 0.567 | 0.18 | 0.566 | 0.179 | 1.302 | 0.411 |
| | | | 3 | 0.555 | 0.165 | 0.56 | 0.165 | 0.558 | 0.165 | 1.381 | 0.465 |
| | | 0.5 | 1 | 1.138 | 0.195 | 1.164 | 0.227 | 1.152 | 0.203 | 2.427 | 0.640 |
| | | | 2 | 0.364 | 0.114 | 0.365 | 0.116 | 0.362 | 0.115 | 0.909 | 0.367 |
| | | | 3 | 0.372 | 0.102 | 0.375 | 0.102 | 0.373 | 0.1 | 0.938 | 0.301 |
| | | 1 | 1 | 0.842 | 0.127 | 0.856 | 0.137 | 0.848 | 0.128 | 1.735 | 0.515 |
| | | | 2 | 0.265 | 0.078 | 0.265 | 0.078 | 0.264 | 0.078 | 0.65 | 0.274 |
| | | | 3 | 0.265 | 0.073 | 0.268 | 0.071 | 0.266 | 0.071 | 0.657 | 0.219 |
| 200 | Gaussian | 0.1 | 1 | 1.264 | 0.173 | 1.263 | 0.173 | 1.263 | 0.173 | 2.543 | 0.528 |
| | | | 2 | 0.42 | 0.126 | 0.418 | 0.124 | 0.418 | 0.124 | 0.908 | 0.262 |
| | | | 3 | 0.401 | 0.111 | 0.399 | 0.11 | 0.399 | 0.11 | 0.934 | 0.255 |
| | | 0.5 | 1 | 0.91 | 0.122 | 0.909 | 0.122 | 0.909 | 0.122 | 1.796 | 0.346 |
| | | | 2 | 0.292 | 0.081 | 0.29 | 0.08 | 0.29 | 0.08 | 0.624 | 0.197 |
| | | | 3 | 0.294 | 0.081 | 0.293 | 0.08 | 0.293 | 0.08 | 0.653 | 0.201 |
| | | 1 | 1 | 0.718 | 0.097 | 0.717 | 0.097 | 0.717 | 0.097 | 1.398 | 0.286 |
| | | | 2 | 0.232 | 0.075 | 0.231 | 0.075 | 0.231 | 0.075 | 0.506 | 0.153 |
| | | | 3 | 0.217 | 0.051 | 0.217 | 0.051 | 0.217 | 0.051 | 0.506 | 0.143 |
| | $t_3$ | 0.1 | 1 | 1.167 | 0.155 | 1.186 | 0.16 | 1.18 | 0.159 | 3.033 | 5.680 |
| | | | 2 | 0.355 | 0.095 | 0.36 | 0.098 | 0.358 | 0.096 | 1.36 | 5.206 |
| | | | 3 | 0.365 | 0.105 | 0.37 | 0.107 | 0.368 | 0.105 | 1.558 | 6.641 |
| | | 0.5 | 1 | 0.761 | 0.104 | 0.77 | 0.105 | 0.767 | 0.105 | 1.867 | 2.577 |
| | | | 2 | 0.229 | 0.053 | 0.233 | 0.056 | 0.231 | 0.054 | 1.048 | 5.154 |
| | | | 3 | 0.237 | 0.063 | 0.24 | 0.062 | 0.238 | 0.063 | 1.101 | 5.291 |
| | | 1 | 1 | 0.585 | 0.077 | 0.588 | 0.078 | 0.586 | 0.078 | 1.183 | 0.244 |
| | | | 2 | 0.173 | 0.047 | 0.173 | 0.044 | 0.172 | 0.043 | 0.411 | 0.139 |
| | | | 3 | 0.184 | 0.044 | 0.184 | 0.043 | 0.183 | 0.043 | 0.411 | 0.125 |
| 500 | Gaussian | 0.1 | 1 | 0.774 | 0.111 | 0.774 | 0.111 | 0.774 | 0.111 | 1.504 | 0.298 |
| | | | 2 | 0.263 | 0.092 | 0.263 | 0.091 | 0.263 | 0.091 | 0.567 | 0.177 |
| | | | 3 | 0.267 | 0.092 | 0.266 | 0.092 | 0.266 | 0.092 | 0.542 | 0.159 |
| | | 0.5 | 1 | 0.556 | 0.078 | 0.556 | 0.078 | 0.556 | 0.078 | 1.064 | 0.192 |
| | | | 2 | 0.184 | 0.058 | 0.184 | 0.057 | 0.184 | 0.057 | 0.384 | 0.120 |
| | | | 3 | 0.187 | 0.056 | 0.187 | 0.056 | 0.187 | 0.056 | 0.395 | 0.116 |
| | | 1 | 1 | 0.442 | 0.064 | 0.441 | 0.064 | 0.441 | 0.064 | 0.824 | 0.140 |
| | | | 2 | 0.141 | 0.047 | 0.141 | 0.047 | 0.141 | 0.047 | 0.308 | 0.096 |
| | | | 3 | 0.149 | 0.042 | 0.149 | 0.042 | 0.149 | 0.042 | 0.302 | 0.079 |
| | $t_3$ | 0.1 | 1 | 0.734 | 0.104 | 0.744 | 0.106 | 0.741 | 0.106 | 1.502 | 0.320 |
| | | | 2 | 0.248 | 0.095 | 0.252 | 0.096 | 0.251 | 0.096 | 0.545 | 0.161 |
| | | | 3 | 0.245 | 0.07 | 0.246 | 0.069 | 0.245 | 0.069 | 0.508 | 0.201 |
| | | 0.5 | 1 | 0.485 | 0.068 | 0.489 | 0.068 | 0.488 | 0.068 | 0.959 | 0.200 |
| | | | 2 | 0.157 | 0.055 | 0.158 | 0.055 | 0.157 | 0.056 | 0.343 | 0.093 |
| | | | 3 | 0.152 | 0.043 | 0.154 | 0.043 | 0.153 | 0.042 | 0.33 | 0.105 |
| | | 1 | 1 | 0.368 | 0.054 | 0.371 | 0.055 | 0.369 | 0.055 | 0.727 | 0.151 |
| | | | 2 | 0.118 | 0.039 | 0.119 | 0.04 | 0.119 | 0.04 | 0.262 | 0.078 |
| | | | 3 | 0.119 | 0.031 | 0.12 | 0.031 | 0.12 | 0.031 | 0.252 | 0.071 |

Table C.12: (T2) Common component estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n-10+1, \ldots, n\}$ ('local') scaled by 1000, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | noTrunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Range | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | All | 2.734 | 0.833 | 2.732 | 0.833 | 2.731 | 0.832 | 2.731 | 0.832 | 3.872 | 1.314 |
| | | | Local | 3.017 | 1.114 | 3.016 | 1.116 | 3.014 | 1.116 | 3.014 | 1.116 | 4.19 | 1.584 |
| | | 0.5 | All | 1.452 | 0.449 | 1.451 | 0.449 | 1.45 | 0.448 | 1.45 | 0.448 | 2.046 | 0.700 |
| | | | Local | 1.758 | 0.925 | 1.758 | 0.925 | 1.758 | 0.925 | 1.758 | 0.925 | 2.38 | 1.098 |
| | | 1 | All | 0.911 | 0.279 | 0.91 | 0.279 | 0.91 | 0.279 | 0.91 | 0.279 | 1.258 | 0.418 |
| | | | Local | 1.078 | 0.813 | 1.078 | 0.813 | 1.078 | 0.813 | 1.078 | 0.813 | 1.431 | 0.887 |
| | $t_3$ | 0.1 | All | 2.552 | 0.799 | 2.636 | 0.843 | 2.648 | 0.852 | 2.644 | 0.85 | 3.98 | 1.580 |
| | | | Local | 3.16 | 1.347 | 3.297 | 1.508 | 3.314 | 1.534 | 3.308 | 1.52 | 4.779 | 2.235 |
| | | 0.5 | All | 1.149 | 0.363 | 1.175 | 0.374 | 1.181 | 0.377 | 1.178 | 0.375 | 1.689 | 0.557 |
| | | | Local | 1.788 | 1.246 | 1.835 | 1.272 | 1.843 | 1.277 | 1.838 | 1.274 | 2.4 | 1.368 |
| | | 1 | All | 0.669 | 0.202 | 0.681 | 0.208 | 0.684 | 0.208 | 0.682 | 0.208 | 0.954 | 0.315 |
| | | | Local | 1.1 | 1.043 | 1.12 | 1.045 | 1.123 | 1.045 | 1.121 | 1.045 | 1.4 | 1.092 |
| 200 | Gaussian | 0.1 | All | 2.684 | 0.714 | 2.683 | 0.713 | 2.683 | 0.713 | 2.683 | 0.713 | 3.263 | 0.885 |
| | | | Local | 2.886 | 1.158 | 2.884 | 1.158 | 2.883 | 1.158 | 2.883 | 1.158 | 3.472 | 1.291 |
| | | 0.5 | All | 1.465 | 0.391 | 1.465 | 0.391 | 1.465 | 0.391 | 1.465 | 0.391 | 1.743 | 0.475 |
| | | | Local | 1.844 | 1.174 | 1.844 | 1.174 | 1.844 | 1.173 | 1.844 | 1.173 | 2.128 | 1.228 |
| | | 1 | All | 0.938 | 0.247 | 0.938 | 0.247 | 0.938 | 0.247 | 0.938 | 0.247 | 1.11 | 0.300 |
| | | | Local | 1.22 | 0.929 | 1.22 | 0.929 | 1.22 | 0.929 | 1.22 | 0.929 | 1.397 | 0.961 |
| | $t_3$ | 0.1 | All | 2.209 | 0.555 | 2.279 | 0.584 | 2.285 | 0.586 | 2.283 | 0.585 | 12.353 | 95.291 |
| | | | Local | 2.849 | 1.459 | 3.014 | 1.679 | 3.021 | 1.69 | 3.018 | 1.685 | 10.568 | 69.896 |
| | | 0.5 | All | 1.025 | 0.263 | 1.05 | 0.274 | 1.051 | 0.274 | 1.051 | 0.274 | 6.444 | 51.611 |
| | | | Local | 1.436 | 1.032 | 1.49 | 1.076 | 1.491 | 1.076 | 1.491 | 1.076 | 8.338 | 66.068 |
| | | 1 | All | 0.62 | 0.16 | 0.63 | 0.165 | 0.631 | 0.165 | 0.63 | 0.165 | 0.755 | 0.210 |
| | | | Local | 0.77 | 0.378 | 0.794 | 0.39 | 0.794 | 0.39 | 0.794 | 0.39 | 0.919 | 0.412 |
| 500 | Gaussian | 0.1 | All | 2.453 | 0.716 | 2.453 | 0.716 | 2.453 | 0.716 | 2.453 | 0.716 | 2.654 | 0.781 |
| | | | Local | 2.65 | 1.021 | 2.65 | 1.021 | 2.65 | 1.021 | 2.65 | 1.021 | 2.853 | 1.057 |
| | | 0.5 | All | 1.352 | 0.382 | 1.352 | 0.382 | 1.352 | 0.382 | 1.352 | 0.382 | 1.449 | 0.407 |
| | | | Local | 1.693 | 0.947 | 1.693 | 0.947 | 1.693 | 0.947 | 1.693 | 0.947 | 1.793 | 0.956 |
| | | 1 | All | 0.864 | 0.248 | 0.864 | 0.248 | 0.864 | 0.248 | 0.864 | 0.248 | 0.923 | 0.264 |
| | | | Local | 1.158 | 0.862 | 1.158 | 0.862 | 1.158 | 0.862 | 1.158 | 0.862 | 1.217 | 0.868 |
| | $t_3$ | 0.1 | All | 2.158 | 0.646 | 2.224 | 0.671 | 2.226 | 0.672 | 2.225 | 0.672 | 2.425 | 0.730 |
| | | | Local | 3.015 | 1.523 | 3.132 | 1.717 | 3.134 | 1.72 | 3.134 | 1.718 | 3.344 | 1.762 |
| | | 0.5 | All | 1.02 | 0.316 | 1.043 | 0.324 | 1.043 | 0.324 | 1.043 | 0.324 | 1.124 | 0.348 |
| | | | Local | 1.842 | 1.528 | 1.889 | 1.646 | 1.889 | 1.647 | 1.889 | 1.646 | 1.973 | 1.658 |
| | | 1 | All | 0.614 | 0.191 | 0.625 | 0.194 | 0.625 | 0.194 | 0.625 | 0.194 | 0.672 | 0.208 |
| | | | Local | 1.021 | 0.988 | 1.079 | 1.289 | 1.08 | 1.29 | 1.079 | 1.289 | 1.128 | 1.298 |

Table C.13: (T3) Loading estimation errors of Trunc, iPE, RTFA and PreAve measured as in (C.2) for each mode scaled by 100, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | 1 | 0.5 | 0.088 | 0.5 | 0.089 | 0.5 | 0.089 | 1.196 | 0.253 |
| | | | 2 | 0.618 | 0.081 | 0.618 | 0.081 | 0.618 | 0.081 | 1.467 | 0.238 |
| | | | 3 | 0.715 | 0.09 | 0.715 | 0.091 | 0.715 | 0.091 | 1.609 | 0.275 |
| | | 0.5 | 1 | 0.359 | 0.06 | 0.359 | 0.06 | 0.359 | 0.06 | 0.853 | 0.190 |
| | | | 2 | 0.45 | 0.068 | 0.449 | 0.068 | 0.449 | 0.068 | 1.07 | 0.235 |
| | | | 3 | 0.513 | 0.068 | 0.512 | 0.068 | 0.512 | 0.068 | 1.186 | 0.273 |
| | | 1 | 1 | 0.282 | 0.047 | 0.281 | 0.047 | 0.281 | 0.047 | 0.681 | 0.172 |
| | | | 2 | 0.341 | 0.046 | 0.341 | 0.046 | 0.341 | 0.046 | 0.829 | 0.198 |
| | | | 3 | 0.396 | 0.049 | 0.396 | 0.049 | 0.396 | 0.049 | 0.902 | 0.184 |
| | $t_3$ | 0.1 | 1 | 0.498 | 0.087 | 0.509 | 0.094 | 0.504 | 0.087 | 1.249 | 0.311 |
| | | | 2 | 0.614 | 0.096 | 0.625 | 0.099 | 0.622 | 0.097 | 1.503 | 0.364 |
| | | | 3 | 0.695 | 0.094 | 0.707 | 0.097 | 0.704 | 0.095 | 1.675 | 0.290 |
| | | 0.5 | 1 | 0.321 | 0.055 | 0.329 | 0.057 | 0.326 | 0.057 | 0.895 | 0.271 |
| | | | 2 | 0.397 | 0.063 | 0.404 | 0.063 | 0.402 | 0.063 | 1.025 | 0.292 |
| | | | 3 | 0.449 | 0.064 | 0.455 | 0.064 | 0.453 | 0.064 | 1.117 | 0.287 |
| | | 1 | 1 | 0.238 | 0.037 | 0.241 | 0.038 | 0.24 | 0.037 | 0.599 | 0.176 |
| | | | 2 | 0.292 | 0.046 | 0.296 | 0.047 | 0.294 | 0.046 | 0.75 | 0.219 |
| | | | 3 | 0.333 | 0.04 | 0.337 | 0.04 | 0.335 | 0.04 | 0.828 | 0.197 |
| 200 | Gaussian | 0.1 | 1 | 0.352 | 0.06 | 0.351 | 0.06 | 0.351 | 0.06 | 0.823 | 0.144 |
| | | | 2 | 0.429 | 0.058 | 0.429 | 0.057 | 0.429 | 0.057 | 0.995 | 0.177 |
| | | | 3 | 0.505 | 0.061 | 0.505 | 0.061 | 0.505 | 0.061 | 1.138 | 0.211 |
| | | 0.5 | 1 | 0.25 | 0.038 | 0.25 | 0.038 | 0.25 | 0.038 | 0.585 | 0.117 |
| | | | 2 | 0.309 | 0.041 | 0.309 | 0.041 | 0.309 | 0.041 | 0.708 | 0.129 |
| | | | 3 | 0.366 | 0.048 | 0.365 | 0.048 | 0.365 | 0.048 | 0.838 | 0.149 |
| | | 1 | 1 | 0.202 | 0.029 | 0.202 | 0.029 | 0.202 | 0.029 | 0.471 | 0.099 |
| | | | 2 | 0.242 | 0.035 | 0.242 | 0.035 | 0.242 | 0.035 | 0.555 | 0.113 |
| | | | 3 | 0.285 | 0.036 | 0.285 | 0.036 | 0.285 | 0.036 | 0.64 | 0.116 |
| | $t_3$ | 0.1 | 1 | 0.332 | 0.054 | 0.336 | 0.055 | 0.335 | 0.055 | 0.826 | 0.219 |
| | | | 2 | 0.401 | 0.059 | 0.406 | 0.059 | 0.405 | 0.06 | 0.975 | 0.190 |
| | | | 3 | 0.471 | 0.064 | 0.479 | 0.065 | 0.477 | 0.065 | 1.135 | 0.230 |
| | | 0.5 | 1 | 0.217 | 0.037 | 0.219 | 0.038 | 0.219 | 0.038 | 0.563 | 0.183 |
| | | | 2 | 0.268 | 0.039 | 0.27 | 0.039 | 0.27 | 0.039 | 0.651 | 0.148 |
| | | | 3 | 0.308 | 0.044 | 0.312 | 0.044 | 0.31 | 0.045 | 0.73 | 0.166 |
| | | 1 | 1 | 0.169 | 0.028 | 0.171 | 0.029 | 0.17 | 0.029 | 0.397 | 0.098 |
| | | | 2 | 0.202 | 0.027 | 0.203 | 0.027 | 0.203 | 0.027 | 0.479 | 0.096 |
| | | | 3 | 0.237 | 0.033 | 0.238 | 0.033 | 0.238 | 0.033 | 0.548 | 0.100 |
| 500 | Gaussian | 0.1 | 1 | 0.216 | 0.029 | 0.216 | 0.029 | 0.216 | 0.029 | 0.502 | 0.082 |
| | | | 2 | 0.27 | 0.036 | 0.27 | 0.036 | 0.27 | 0.036 | 0.61 | 0.091 |
| | | | 3 | 0.32 | 0.036 | 0.319 | 0.036 | 0.319 | 0.036 | 0.703 | 0.086 |
| | | 0.5 | 1 | 0.155 | 0.022 | 0.155 | 0.022 | 0.155 | 0.022 | 0.363 | 0.062 |
| | | | 2 | 0.196 | 0.026 | 0.195 | 0.026 | 0.195 | 0.026 | 0.445 | 0.075 |
| | | | 3 | 0.229 | 0.027 | 0.229 | 0.027 | 0.229 | 0.027 | 0.508 | 0.066 |
| | | 1 | 1 | 0.126 | 0.019 | 0.126 | 0.019 | 0.126 | 0.019 | 0.292 | 0.054 |
| | | | 2 | 0.156 | 0.021 | 0.156 | 0.021 | 0.156 | 0.021 | 0.355 | 0.058 |
| | | | 3 | 0.18 | 0.022 | 0.18 | 0.022 | 0.18 | 0.022 | 0.403 | 0.057 |
| | $t_3$ | 0.1 | 1 | 0.205 | 0.031 | 0.208 | 0.031 | 0.207 | 0.031 | 0.486 | 0.083 |
| | | | 2 | 0.242 | 0.027 | 0.245 | 0.029 | 0.244 | 0.028 | 0.573 | 0.089 |
| | | | 3 | 0.284 | 0.035 | 0.288 | 0.036 | 0.288 | 0.036 | 0.642 | 0.116 |
| | | 0.5 | 1 | 0.135 | 0.019 | 0.137 | 0.019 | 0.136 | 0.019 | 0.316 | 0.053 |
| | | | 2 | 0.164 | 0.02 | 0.166 | 0.02 | 0.165 | 0.02 | 0.377 | 0.068 |
| | | | 3 | 0.188 | 0.021 | 0.19 | 0.022 | 0.19 | 0.022 | 0.426 | 0.069 |
| | | 1 | 1 | 0.103 | 0.015 | 0.104 | 0.015 | 0.103 | 0.015 | 0.242 | 0.043 |
| | | | 2 | 0.125 | 0.015 | 0.126 | 0.015 | 0.125 | 0.015 | 0.289 | 0.045 |
| | | | 3 | 0.146 | 0.016 | 0.147 | 0.017 | 0.146 | 0.017 | 0.324 | 0.051 |

Table C.14: (T3) Common component estimation errors of Trunc, noTrunc, iPE, RTFA and PreAve measured as in (C.3) with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n-10+1, \ldots, n\}$ ('local') scaled by 1000, over varying $n \in \{100, 200, 500\}$, the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentages of outliers in the factors under (O2) ($\varrho \in \{0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations for each setting.

| | | | | Trunc | | noTrunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Dist | % | Range | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 100 | Gaussian | 0.1 | All | 1.082 | 0.221 | 1.082 | 0.221 | 1.082 | 0.221 | 1.082 | 0.221 | 1.527 | 0.340 |
| | | | Local | 1.169 | 0.389 | 1.169 | 0.389 | 1.169 | 0.389 | 1.169 | 0.389 | 1.622 | 0.500 |
| | | 0.5 | All | 0.579 | 0.115 | 0.578 | 0.114 | 0.578 | 0.114 | 0.578 | 0.114 | 0.816 | 0.189 |
| | | | Local | 0.773 | 0.414 | 0.772 | 0.414 | 0.772 | 0.414 | 0.772 | 0.414 | 1.012 | 0.458 |
| | | 1 | All | 0.359 | 0.068 | 0.359 | 0.068 | 0.359 | 0.068 | 0.359 | 0.068 | 0.498 | 0.103 |
| | | | Local | 0.472 | 0.297 | 0.472 | 0.297 | 0.471 | 0.297 | 0.471 | 0.297 | 0.613 | 0.326 |
| | $t_3$ | 0.1 | All | 0.983 | 0.226 | 1.018 | 0.239 | 1.021 | 0.242 | 1.02 | 0.239 | 1.475 | 0.364 |
| | | | Local | 1.172 | 0.523 | 1.204 | 0.539 | 1.208 | 0.538 | 1.207 | 0.538 | 1.696 | 0.671 |
| | | 0.5 | All | 0.441 | 0.096 | 0.453 | 0.1 | 0.454 | 0.101 | 0.453 | 0.1 | 0.681 | 0.180 |
| | | | Local | 0.619 | 0.492 | 0.63 | 0.497 | 0.631 | 0.497 | 0.631 | 0.497 | 0.863 | 0.538 |
| | | 1 | All | 0.258 | 0.055 | 0.264 | 0.058 | 0.265 | 0.059 | 0.264 | 0.058 | 0.383 | 0.098 |
| | | | Local | 0.325 | 0.251 | 0.329 | 0.252 | 0.33 | 0.252 | 0.33 | 0.252 | 0.458 | 0.290 |
| 200 | Gaussian | 0.1 | All | 1.032 | 0.223 | 1.032 | 0.223 | 1.032 | 0.223 | 1.032 | 0.223 | 1.253 | 0.285 |
| | | | Local | 1.116 | 0.413 | 1.116 | 0.413 | 1.116 | 0.413 | 1.116 | 0.413 | 1.339 | 0.463 |
| | | 0.5 | All | 0.565 | 0.123 | 0.564 | 0.123 | 0.564 | 0.123 | 0.564 | 0.123 | 0.679 | 0.153 |
| | | | Local | 0.699 | 0.439 | 0.699 | 0.439 | 0.699 | 0.439 | 0.699 | 0.439 | 0.812 | 0.460 |
| | | 1 | All | 0.361 | 0.079 | 0.361 | 0.079 | 0.361 | 0.079 | 0.361 | 0.079 | 0.431 | 0.102 |
| | | | Local | 0.451 | 0.321 | 0.451 | 0.321 | 0.451 | 0.322 | 0.451 | 0.322 | 0.521 | 0.340 |
| | $t_3$ | 0.1 | All | 0.9 | 0.189 | 0.928 | 0.199 | 0.929 | 0.2 | 0.929 | 0.2 | 1.142 | 0.253 |
| | | | Local | 1.104 | 0.473 | 1.138 | 0.487 | 1.139 | 0.487 | 1.139 | 0.487 | 1.358 | 0.516 |
| | | 0.5 | All | 0.416 | 0.088 | 0.425 | 0.091 | 0.426 | 0.091 | 0.426 | 0.091 | 0.52 | 0.116 |
| | | | Local | 0.576 | 0.401 | 0.588 | 0.406 | 0.588 | 0.406 | 0.588 | 0.406 | 0.684 | 0.419 |
| | | 1 | All | 0.252 | 0.054 | 0.257 | 0.056 | 0.257 | 0.056 | 0.257 | 0.056 | 0.308 | 0.068 |
| | | | Local | 0.314 | 0.239 | 0.319 | 0.241 | 0.32 | 0.241 | 0.32 | 0.241 | 0.372 | 0.245 |
| 500 | Gaussian | 0.1 | All | 0.984 | 0.176 | 0.984 | 0.176 | 0.984 | 0.176 | 0.984 | 0.176 | 1.066 | 0.189 |
| | | | Local | 1.146 | 0.334 | 1.146 | 0.334 | 1.146 | 0.334 | 1.146 | 0.334 | 1.229 | 0.344 |
| | | 0.5 | All | 0.545 | 0.097 | 0.545 | 0.097 | 0.545 | 0.097 | 0.545 | 0.097 | 0.588 | 0.105 |
| | | | Local | 0.73 | 0.376 | 0.73 | 0.376 | 0.73 | 0.376 | 0.73 | 0.376 | 0.774 | 0.379 |
| | | 1 | All | 0.349 | 0.062 | 0.349 | 0.062 | 0.349 | 0.062 | 0.349 | 0.062 | 0.377 | 0.067 |
| | | | Local | 0.472 | 0.301 | 0.472 | 0.301 | 0.472 | 0.301 | 0.472 | 0.301 | 0.5 | 0.303 |
| | $t_3$ | 0.1 | All | 0.829 | 0.151 | 0.854 | 0.156 | 0.854 | 0.156 | 0.854 | 0.156 | 0.925 | 0.171 |
| | | | Local | 1.104 | 0.521 | 1.136 | 0.54 | 1.136 | 0.54 | 1.136 | 0.54 | 1.208 | 0.546 |
| | | 0.5 | All | 0.391 | 0.069 | 0.4 | 0.07 | 0.4 | 0.07 | 0.4 | 0.07 | 0.431 | 0.076 |
| | | | Local | 0.631 | 0.465 | 0.648 | 0.489 | 0.648 | 0.489 | 0.648 | 0.489 | 0.679 | 0.492 |
| | | 1 | All | 0.235 | 0.041 | 0.239 | 0.042 | 0.239 | 0.042 | 0.239 | 0.042 | 0.257 | 0.046 |
| | | | Local | 0.354 | 0.312 | 0.36 | 0.316 | 0.36 | 0.316 | 0.36 | 0.316 | 0.378 | 0.316 |

### C.1.3 Factor number estimation

See Tables C.15–C.23 for the results from factor number estimation obtained under (T1)–(T3), with outliers introduced to the idiosyncratic component under (O1).

Table C.15: (T1) with $n = 100$. Factor number estimation results from Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, the distribution for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentage of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0, 0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations per each setting.

| | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Gaussian | 0 | 1 | 2.55 | 0.657 | 2.92 | 0.307 | 2.92 | 0.307 | 2.81 | 0.394 |
| | | 2 | 2.54 | 0.688 | 2.97 | 0.223 | 2.97 | 0.223 | 2.81 | 0.394 |
| | | 3 | 2.59 | 0.683 | 2.96 | 0.197 | 2.96 | 0.197 | 2.86 | 0.349 |
| | 0.1 | 1 | 2.51 | 0.689 | 2.89 | 0.345 | 2.89 | 0.345 | 2.82 | 0.386 |
| | | 2 | 2.54 | 0.688 | 2.95 | 0.261 | 2.96 | 0.243 | 2.81 | 0.394 |
| | | 3 | 2.59 | 0.683 | 2.95 | 0.219 | 2.95 | 0.219 | 2.88 | 0.327 |
| | 0.5 | 1 | 2.53 | 0.674 | 2.78 | 0.484 | 2.77 | 0.489 | 2.78 | 0.416 |
| | | 2 | 2.53 | 0.688 | 2.88 | 0.409 | 2.87 | 0.418 | 2.8 | 0.402 |
| | | 3 | 2.59 | 0.683 | 2.89 | 0.373 | 2.89 | 0.373 | 2.84 | 0.368 |
| | 1 | 1 | 2.49 | 0.703 | 2.69 | 0.563 | 2.7 | 0.56 | 2.67 | 0.493 |
| | | 2 | 2.51 | 0.689 | 2.75 | 0.539 | 2.76 | 0.534 | 2.8 | 0.402 |
| | | 3 | 2.57 | 0.7 | 2.8 | 0.492 | 2.8 | 0.492 | 2.76 | 0.452 |
| t | 0 | 1 | 2.56 | 0.729 | 3 | 0.246 | 2.99 | 0.225 | 2.8 | 0.426 |
| | | 2 | 2.53 | 0.703 | 2.96 | 0.425 | 2.91 | 0.404 | 2.68 | 0.490 |
| | | 3 | 2.49 | 0.759 | 3 | 0.284 | 2.97 | 0.223 | 2.74 | 0.441 |
| | 0.1 | 1 | 2.53 | 0.745 | 2.96 | 0.315 | 2.96 | 0.315 | 2.83 | 0.403 |
| | | 2 | 2.51 | 0.718 | 2.91 | 0.452 | 2.91 | 0.404 | 2.69 | 0.506 |
| | | 3 | 2.47 | 0.758 | 2.98 | 0.284 | 2.96 | 0.243 | 2.71 | 0.456 |
| | 0.5 | 1 | 2.49 | 0.772 | 2.75 | 0.657 | 2.79 | 0.591 | 2.73 | 0.489 |
| | | 2 | 2.48 | 0.731 | 2.77 | 0.601 | 2.81 | 0.526 | 2.61 | 0.549 |
| | | 3 | 2.45 | 0.77 | 2.82 | 0.557 | 2.81 | 0.545 | 2.69 | 0.486 |
| | 1 | 1 | 2.43 | 0.807 | 2.59 | 0.726 | 2.6 | 0.725 | 2.68 | 0.530 |
| | | 2 | 2.48 | 0.731 | 2.64 | 0.659 | 2.64 | 0.674 | 2.53 | 0.594 |
| | | 3 | 2.43 | 0.782 | 2.71 | 0.64 | 2.71 | 0.64 | 2.62 | 0.528 |

Table C.16: (T1) with $n = 200$.

| | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Gaussian | 0 | 1 | 2.61 | 0.665 | 2.92 | 0.273 | 2.92 | 0.273 | 2.89 | 0.314 |
| | | 2 | 2.55 | 0.73 | 2.95 | 0.261 | 2.95 | 0.261 | 2.9 | 0.302 |
| | | 3 | 2.57 | 0.685 | 2.95 | 0.261 | 2.95 | 0.261 | 2.84 | 0.368 |
| | 0.1 | 1 | 2.59 | 0.683 | 2.92 | 0.273 | 2.92 | 0.273 | 2.88 | 0.327 |
| | | 2 | 2.53 | 0.745 | 2.95 | 0.261 | 2.95 | 0.261 | 2.91 | 0.288 |
| | | 3 | 2.55 | 0.702 | 2.91 | 0.351 | 2.91 | 0.351 | 2.84 | 0.368 |
| | 0.5 | 1 | 2.59 | 0.683 | 2.85 | 0.386 | 2.87 | 0.338 | 2.9 | 0.302 |
| | | 2 | 2.53 | 0.731 | 2.87 | 0.442 | 2.88 | 0.433 | 2.87 | 0.338 |
| | | 3 | 2.55 | 0.702 | 2.85 | 0.435 | 2.85 | 0.435 | 2.83 | 0.378 |
| | 1 | 1 | 2.54 | 0.731 | 2.76 | 0.534 | 2.79 | 0.498 | 2.86 | 0.349 |
| | | 2 | 2.52 | 0.731 | 2.79 | 0.537 | 2.8 | 0.532 | 2.84 | 0.368 |
| | | 3 | 2.55 | 0.702 | 2.79 | 0.498 | 2.79 | 0.498 | 2.81 | 0.394 |
| t | 0 | 1 | 2.66 | 0.623 | 2.91 | 0.404 | 2.91 | 0.379 | 2.75 | 0.435 |
| | | 2 | 2.5 | 0.732 | 2.97 | 0.332 | 2.95 | 0.297 | 2.72 | 0.473 |
| | | 3 | 2.4 | 0.778 | 2.95 | 0.359 | 2.95 | 0.33 | 2.63 | 0.506 |
| | 0.1 | 1 | 2.63 | 0.63 | 2.89 | 0.399 | 2.88 | 0.433 | 2.71 | 0.478 |
| | | 2 | 2.43 | 0.756 | 2.95 | 0.33 | 2.94 | 0.312 | 2.67 | 0.514 |
| | | 3 | 2.41 | 0.793 | 2.91 | 0.429 | 2.92 | 0.367 | 2.63 | 0.506 |
| | 0.5 | 1 | 2.61 | 0.65 | 2.79 | 0.537 | 2.79 | 0.537 | 2.71 | 0.478 |
| | | 2 | 2.44 | 0.756 | 2.8 | 0.512 | 2.85 | 0.479 | 2.66 | 0.517 |
| | | 3 | 2.35 | 0.796 | 2.78 | 0.561 | 2.83 | 0.493 | 2.61 | 0.530 |
| | 1 | 1 | 2.57 | 0.671 | 2.62 | 0.708 | 2.61 | 0.695 | 2.69 | 0.486 |
| | | 2 | 2.44 | 0.743 | 2.74 | 0.597 | 2.7 | 0.595 | 2.66 | 0.497 |
| | | 3 | 2.34 | 0.794 | 2.56 | 0.715 | 2.59 | 0.698 | 2.56 | 0.574 |

Table C.17: (T1) with $n = 500$.

| Dist | % | Mode | Trunc Mean | Trunc SD | iPE Mean | iPE SD | RTFA Mean | RTFA SD | PreAve Mean | PreAve SD |
|---|---|---|---|---|---|---|---|---|---|---|
| Gaussian | 0 | 1 | 2.56 | 0.671 | 2.91 | 0.288 | 2.91 | 0.288 | 2.84 | 0.368 |
| | | 2 | 2.56 | 0.641 | 2.91 | 0.321 | 2.91 | 0.321 | 2.82 | 0.386 |
| | | 3 | 2.57 | 0.714 | 2.97 | 0.223 | 2.97 | 0.223 | 2.88 | 0.327 |
| | 0.1 | 1 | 2.53 | 0.688 | 2.85 | 0.359 | 2.85 | 0.359 | 2.85 | 0.359 |
| | | 2 | 2.55 | 0.642 | 2.9 | 0.333 | 2.9 | 0.333 | 2.82 | 0.386 |
| | | 3 | 2.57 | 0.714 | 2.96 | 0.243 | 2.96 | 0.243 | 2.88 | 0.327 |
| | 0.5 | 1 | 2.49 | 0.718 | 2.75 | 0.539 | 2.75 | 0.539 | 2.84 | 0.368 |
| | | 2 | 2.55 | 0.642 | 2.81 | 0.443 | 2.83 | 0.428 | 2.81 | 0.394 |
| | | 3 | 2.56 | 0.715 | 2.91 | 0.351 | 2.91 | 0.351 | 2.88 | 0.327 |
| | 1 | 1 | 2.48 | 0.717 | 2.67 | 0.604 | 2.67 | 0.604 | 2.81 | 0.394 |
| | | 2 | 2.55 | 0.642 | 2.73 | 0.51 | 2.73 | 0.51 | 2.8 | 0.402 |
| | | 3 | 2.54 | 0.731 | 2.84 | 0.465 | 2.86 | 0.427 | 2.87 | 0.338 |
| t | 0 | 1 | 2.61 | 0.68 | 2.86 | 0.427 | 2.84 | 0.443 | 2.8 | 0.426 |
| | | 2 | 2.55 | 0.716 | 2.91 | 0.379 | 2.91 | 0.379 | 2.86 | 0.377 |
| | | 3 | 2.63 | 0.646 | 2.89 | 0.399 | 2.88 | 0.409 | 2.83 | 0.403 |
| | 0.1 | 1 | 2.61 | 0.68 | 2.83 | 0.451 | 2.83 | 0.451 | 2.79 | 0.433 |
| | | 2 | 2.55 | 0.716 | 2.86 | 0.472 | 2.86 | 0.472 | 2.87 | 0.367 |
| | | 3 | 2.58 | 0.684 | 2.85 | 0.458 | 2.86 | 0.45 | 2.78 | 0.440 |
| | 0.5 | 1 | 2.59 | 0.653 | 2.76 | 0.534 | 2.75 | 0.539 | 2.78 | 0.462 |
| | | 2 | 2.55 | 0.716 | 2.75 | 0.575 | 2.73 | 0.601 | 2.83 | 0.403 |
| | | 3 | 2.58 | 0.699 | 2.78 | 0.543 | 2.79 | 0.518 | 2.77 | 0.468 |
| | 1 | 1 | 2.54 | 0.702 | 2.56 | 0.729 | 2.57 | 0.714 | 2.73 | 0.489 |
| | | 2 | 2.57 | 0.7 | 2.55 | 0.744 | 2.59 | 0.712 | 2.8 | 0.426 |
| | | 3 | 2.53 | 0.731 | 2.7 | 0.644 | 2.7 | 0.644 | 2.78 | 0.440 |

Table C.18: (T2) with $n = 100$. Factor number estimation results from Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, the distribution for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentage of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0, 0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations per each setting.

| | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Gaussian | 0 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.02 | 0.141 |
| | | 2 | 2.73 | 0.548 | 2.99 | 0.1 | 2.99 | 0.1 | 2.88 | 0.327 |
| | | 3 | 2.81 | 0.465 | 3 | 0 | 3 | 0 | 2.89 | 0.314 |
| | 0.1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.02 | 0.141 |
| | | 2 | 2.73 | 0.548 | 2.99 | 0.1 | 2.99 | 0.1 | 2.87 | 0.338 |
| | | 3 | 2.81 | 0.465 | 3 | 0 | 3 | 0 | 2.87 | 0.338 |
| | 0.5 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.71 | 0.574 | 2.98 | 0.141 | 2.98 | 0.141 | 2.87 | 0.338 |
| | | 3 | 2.8 | 0.471 | 3 | 0 | 3 | 0 | 2.87 | 0.338 |
| | 1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.7 | 0.577 | 2.98 | 0.141 | 2.98 | 0.141 | 2.87 | 0.338 |
| | | 3 | 2.8 | 0.471 | 3 | 0 | 3 | 0 | 2.83 | 0.378 |
| t | 0 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.733 | 0.546 | 3 | 0 | 3 | 0 | 2.713 | 0.476 |
| | | 3 | 2.683 | 0.615 | 3 | 0 | 3 | 0 | 2.713 | 0.476 |
| | 0.1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.713 | 0.554 | 3 | 0 | 3 | 0 | 2.693 | 0.485 |
| | | 3 | 2.683 | 0.615 | 3 | 0 | 3 | 0 | 2.693 | 0.485 |
| | 0.5 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.693 | 0.579 | 2.98 | 0.14 | 2.98 | 0.14 | 2.693 | 0.485 |
| | | 3 | 2.663 | 0.621 | 2.99 | 0.1 | 2.99 | 0.1 | 2.663 | 0.496 |
| | 1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.455 | 0.933 |
| | | 2 | 2.693 | 0.579 | 2.96 | 0.196 | 2.96 | 0.196 | 2.683 | 0.488 |
| | | 3 | 2.634 | 0.644 | 2.96 | 0.196 | 2.96 | 0.196 | 2.673 | 0.492 |

Table C.19: (T2) with $n = 200$.

| Dist | % | Mode | Trunc Mean | Trunc SD | iPE Mean | iPE SD | RTFA Mean | RTFA SD | PreAve Mean | PreAve SD |
|---|---|---|---|---|---|---|---|---|---|---|
| Gaussian | 0 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.79 | 0.478 | 2.99 | 0.1 | 2.99 | 0.1 | 2.89 | 0.314 |
| | | 3 | 2.72 | 0.57 | 3 | 0 | 3 | 0 | 2.89 | 0.314 |
| | 0.1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.79 | 0.478 | 2.99 | 0.1 | 2.99 | 0.1 | 2.89 | 0.314 |
| | | 3 | 2.71 | 0.591 | 3 | 0 | 3 | 0 | 2.89 | 0.314 |
| | 0.5 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.78 | 0.484 | 2.99 | 0.1 | 2.99 | 0.1 | 2.88 | 0.327 |
| | | 3 | 2.7 | 0.595 | 3 | 0 | 3 | 0 | 2.86 | 0.349 |
| | 1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.77 | 0.489 | 2.97 | 0.171 | 2.97 | 0.171 | 2.87 | 0.338 |
| | | 3 | 2.69 | 0.598 | 3 | 0 | 3 | 0 | 2.85 | 0.359 |
| t | 0 | 1 | 3 | 0 | 3.02 | 0.141 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.798 | 0.494 | 3.01 | 0.175 | 2.99 | 0.101 | 2.828 | 0.379 |
| | | 3 | 2.758 | 0.454 | 3.02 | 0.141 | 3 | 0 | 2.737 | 0.442 |
| | 0.1 | 1 | 3 | 0 | 3.02 | 0.141 | 3 | 0 | 3.01 | 0.101 |
| | | 2 | 2.798 | 0.494 | 3.01 | 0.175 | 2.99 | 0.101 | 2.798 | 0.404 |
| | | 3 | 2.758 | 0.454 | 3.02 | 0.141 | 3 | 0 | 2.727 | 0.448 |
| | 0.5 | 1 | 3 | 0 | 3.01 | 0.101 | 3 | 0 | 3.091 | 0.380 |
| | | 2 | 2.778 | 0.526 | 2.97 | 0.224 | 2.96 | 0.198 | 2.798 | 0.428 |
| | | 3 | 2.747 | 0.459 | 3.01 | 0.101 | 3 | 0 | 2.717 | 0.453 |
| | 1 | 1 | 3.01 | 0.101 | 3.02 | 0.141 | 3.01 | 0.101 | 4.596 | 1.911 |
| | | 2 | 2.768 | 0.531 | 2.97 | 0.266 | 2.96 | 0.244 | 2.788 | 0.435 |
| | | 3 | 2.747 | 0.459 | 3 | 0 | 3 | 0 | 2.667 | 0.535 |

Table C.20: (T2) with $n = 500$.

| Dist | % | Mode | Trunc Mean | Trunc SD | iPE Mean | iPE SD | RTFA Mean | RTFA SD | PreAve Mean | PreAve SD |
|---|---|---|---|---|---|---|---|---|---|---|
| Gaussian | 0 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.01 | 0.100 |
| | | 2 | 2.8 | 0.426 | 3 | 0 | 3 | 0 | 2.87 | 0.338 |
| | | 3 | 2.66 | 0.623 | 3 | 0 | 3 | 0 | 2.88 | 0.327 |
| | 0.1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.01 | 0.100 |
| | | 2 | 2.8 | 0.426 | 3 | 0 | 3 | 0 | 2.88 | 0.327 |
| | | 3 | 2.66 | 0.623 | 3 | 0 | 3 | 0 | 2.89 | 0.314 |
| | 0.5 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.78 | 0.462 | 2.99 | 0.1 | 2.99 | 0.1 | 2.87 | 0.338 |
| | | 3 | 2.64 | 0.644 | 3 | 0 | 3 | 0 | 2.89 | 0.314 |
| | 1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.02 | 0.141 |
| | | 2 | 2.77 | 0.468 | 2.97 | 0.171 | 2.97 | 0.171 | 2.88 | 0.327 |
| | | 3 | 2.64 | 0.644 | 3 | 0 | 3 | 0 | 2.89 | 0.314 |
| t | 0 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 2.69 | 0.563 | 3 | 0 | 3 | 0 | 2.85 | 0.359 |
| | | 3 | 2.86 | 0.427 | 3 | 0 | 3 | 0 | 2.88 | 0.327 |
| | 0.1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.01 | 0.100 |
| | | 2 | 2.69 | 0.563 | 3 | 0 | 3 | 0 | 2.84 | 0.368 |
| | | 3 | 2.85 | 0.435 | 2.99 | 0.1 | 2.99 | 0.1 | 2.88 | 0.327 |
| | 0.5 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3.4 | 0.791 |
| | | 2 | 2.64 | 0.595 | 2.98 | 0.141 | 2.97 | 0.171 | 2.81 | 0.394 |
| | | 3 | 2.79 | 0.537 | 2.99 | 0.1 | 2.99 | 0.1 | 2.89 | 0.314 |
| | 1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 6.62 | 3.296 |
| | | 2 | 2.63 | 0.597 | 2.93 | 0.256 | 2.93 | 0.256 | 2.82 | 0.386 |
| | | 3 | 2.79 | 0.537 | 2.96 | 0.243 | 2.96 | 0.243 | 2.86 | 0.349 |

Table C.21: (T3) with $n = 100$. Factor number estimation results from Trunc, iPE, RTFA and PreAve over varying $n \in \{100, 200, 500\}$, the distribution for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$) and the percentage of outliers in the idiosyncratic component under (O1) ($\varrho \in \{0, 0.1, 0.5, 1\}$). We report the mean and the standard deviation over 100 realisations per each setting.

| | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Gaussian | 0 | 1 | 2.98 | 0.2 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.1 | 1 | 2.97 | 0.223 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.5 | 1 | 2.97 | 0.223 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 1 | 1 | 2.97 | 0.223 | 2.98 | 0.2 | 2.98 | 0.2 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| t | 0 | 1 | 2.99 | 0.1 | 3.03 | 0.223 | 3.03 | 0.223 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3.03 | 0.223 | 3.03 | 0.223 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3.03 | 0.223 | 3.03 | 0.223 | 3 | 0.000 |
| | 0.1 | 1 | 2.97 | 0.223 | 3.02 | 0.2 | 3.02 | 0.2 | 2.97 | 0.171 |
| | | 2 | 2.98 | 0.2 | 3.03 | 0.223 | 3.02 | 0.2 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3.03 | 0.223 | 3.02 | 0.2 | 3 | 0.000 |
| | 0.5 | 1 | 2.99 | 0.301 | 2.99 | 0.301 | 2.99 | 0.301 | 2.98 | 0.141 |
| | | 2 | 2.99 | 0.225 | 3 | 0.284 | 3 | 0.284 | 2.99 | 0.100 |
| | | 3 | 3.02 | 0.2 | 3.02 | 0.2 | 3.02 | 0.2 | 3 | 0.000 |
| | 1 | 1 | 2.96 | 0.243 | 2.92 | 0.394 | 2.94 | 0.343 | 2.99 | 0.100 |
| | | 2 | 2.98 | 0.2 | 2.99 | 0.225 | 3.01 | 0.1 | 3.02 | 0.200 |
| | | 3 | 3 | 0 | 3.01 | 0.1 | 3.01 | 0.1 | 3.01 | 0.100 |

Table C.22: (T3) with $n = 200$.

| | | | Trunc | | iPE | | RTFA | | PreAve | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dist | % | Mode | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Gaussian | 0 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.5 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| t | 0 | 1 | 2.98 | 0.141 | 3 | 0 | 3 | 0 | 2.99 | 0.100 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 2.99 | 0.100 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.1 | 1 | 2.96 | 0.243 | 2.97 | 0.223 | 2.97 | 0.223 | 2.99 | 0.100 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 2.99 | 0.100 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.5 | 1 | 2.96 | 0.243 | 2.93 | 0.326 | 2.93 | 0.326 | 2.99 | 0.100 |
| | | 2 | 3 | 0 | 2.98 | 0.2 | 2.98 | 0.2 | 2.99 | 0.100 |
| | | 3 | 3 | 0 | 2.98 | 0.2 | 2.98 | 0.2 | 3 | 0.000 |
| | 1 | 1 | 2.95 | 0.297 | 2.91 | 0.379 | 2.92 | 0.367 | 2.98 | 0.141 |
| | | 2 | 2.98 | 0.2 | 2.98 | 0.2 | 2.98 | 0.2 | 2.99 | 0.100 |
| | | 3 | 2.98 | 0.2 | 2.98 | 0.2 | 2.98 | 0.2 | 3.02 | 0.141 |

Table C.23: (T3) with $n = 500$.

| Dist | % | Mode | Trunc Mean | Trunc SD | iPE Mean | iPE SD | RTFA Mean | RTFA SD | PreAve Mean | PreAve SD |
|---|---|---|---|---|---|---|---|---|---|---|
| Gaussian | 0 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.5 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 1 | 1 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| t | 0 | 1 | 2.99 | 0.1 | 3 | 0 | 3 | 0 | 2.98 | 0.141 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.1 | 1 | 2.99 | 0.1 | 3 | 0 | 3 | 0 | 2.98 | 0.141 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 0.5 | 1 | 2.98 | 0.141 | 2.96 | 0.197 | 2.96 | 0.197 | 2.98 | 0.141 |
| | | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 0.000 |
| | 1 | 1 | 2.97 | 0.171 | 2.89 | 0.373 | 2.89 | 0.373 | 2.98 | 0.141 |
| | | 2 | 3 | 0 | 2.99 | 0.1 | 2.99 | 0.1 | 3 | 0.000 |
| | | 3 | 3 | 0 | 3 | 0 | 3 | 0 | 3.01 | 0.100 |

### C.1.4 Asymptotic normality

We investigate the asymptotic normality of the loading estimator $\check{\mathbf{\Lambda}}_k^{[2]}$. For this, we generate the tensor time series as described in Section C.1.1 but to avoid the estimation of the rotation matrix $\check{\mathbf{H}}_k^{[2]}$, set $(r_1, r_2, r_3) = (1, 1, 1)$ and scale (now vectors) $\mathbf{\Lambda}_k = (\lambda_{k,11}, \ldots, \lambda_{k,p_k 1})^\top \in \mathbb{R}^{p_k}$, to have $|\mathbf{\Lambda}_k|_2 = 1$ for all $k \in [K]$. Also, to study the validity of the asymptotic normality without being hampered by the difficulties from estimating the long-run variance in the limit in Theorem 3 (ii), particularly in the presence of heavy tails, we set $\phi = \psi = 0$.
For each realisation, we compute

$$Z_{k,i} = \frac{\sqrt{np_{-k}}(\check{\lambda}_{k,i}^{[2]}(\tau) - s_k \lambda_{k,i1})}{(\widehat{\Phi}_i^{(k)}(\tau))^{1/2}} \quad \text{for all} \quad i \in [p_k] \quad \text{and} \quad k \in [K],$$

where $s_k \in \{-1, 1\}$ denotes the sign set to be $\text{median}_{i \in [p_k]} \, \mathsf{sign}(\check{\lambda}_{k,i}^{[2]}(\tau) \cdot \lambda_{k,i1})$. The variance estimator $\widehat{\Phi}_i^{(k)}(\tau)$ is obtained as

$$\widehat{\Phi}_i^{(k)}(\tau) = (\widehat{\mathbf{\Gamma}}_f^{(k)})^{-1} \widehat{\mathsf{Var}} \left[ \left( \mathsf{mat}_k(\mathcal{X}_t^{\mathsf{t}}(\tau))_{i\cdot} - \mathsf{mat}_k(\widehat{\boldsymbol{\chi}}_t^{\mathsf{t}}(\tau))_{i\cdot} \right) \check{\mathbf{D}}_k^{[2]} \right] (\widehat{\mathbf{\Gamma}}_f^{(k)})^{-1},$$

where $\widehat{\mathbf{\Gamma}}_f^{(k)} = n^{-1} \sum_{t \in [n]} \widehat{\mathcal{F}}_t^2$, $\widehat{\boldsymbol{\chi}}_t$ denotes the estimator of the common component and $\widehat{\boldsymbol{\chi}}_t^{\mathsf{t}}(\tau)$ its truncated version, $\check{\mathbf{D}}_k^{[2]} = p_{-k}^{-1/2} \check{\mathbf{\Lambda}}_K^{[2]} \otimes \ldots \otimes \check{\mathbf{\Lambda}}_{k+1}^{[2]} \otimes \check{\mathbf{\Lambda}}_{k-1}^{[2]} \otimes \ldots \otimes \check{\mathbf{\Lambda}}_1^{[2]}$, and $\widehat{\mathsf{Var}}(\cdot)$ denotes the sample variance operator. Repeating the experiments over $B = 100$ realisations, we generate the Q-Q plot of $Z_{b,k,i}$, $i \in [p_k]$, $k \in [K]$, $b \in [B]$, against the standard normal distribution for each setting while varying $(p_1, p_2, p_3)$ under (T1)–(T3), distributions of $\boldsymbol{\xi}_t$ and $\mathcal{F}_t$ (Gaussian and $t_3$) and $n \in \{100, 200, 500\}$, see Figures C.15–C.17.
We observe that as $n$ (within each figure) and $p_k$'s (from (T1) to (T3)) increase, the approximation by the standard normal distribution improves, even when the data are generated from the $t_3$ distribution. While the asymptotic normality of the loading estimator in Theorem 3 (ii) is derived under spatial independence, the mild cross-sectional dependence introduced as described in Appendix C.1.1 does not influence the approximation greatly.

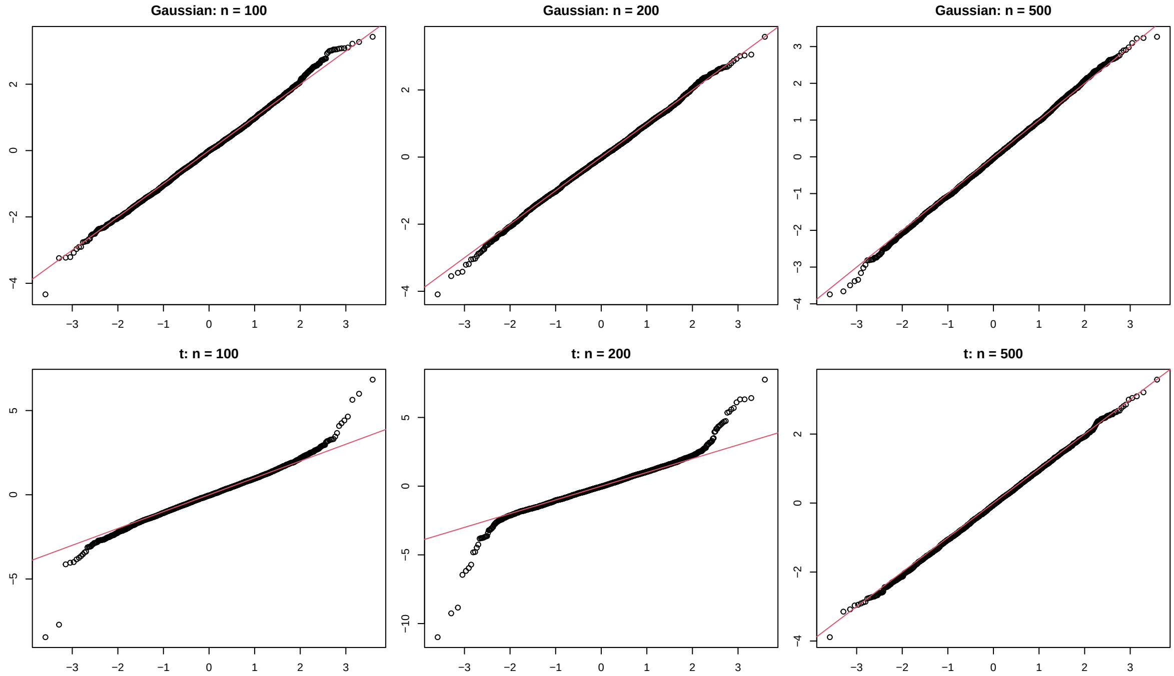


Figure C.15: (T1) Q-Q plot of the sample quantiles of $Z_{b,k,i}$, $i \in [p_k]$, $k \in [K]$, $b \in [B]$ ($y$-axis), against the quantiles from the standard normal distribution ($x$-axis) over varying $n \in \{100, 200, 500\}$ (left to right) and the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$, top to bottom) over $B = 100$ realisations per setting. In each plot, the $y = x$ line is given in red.

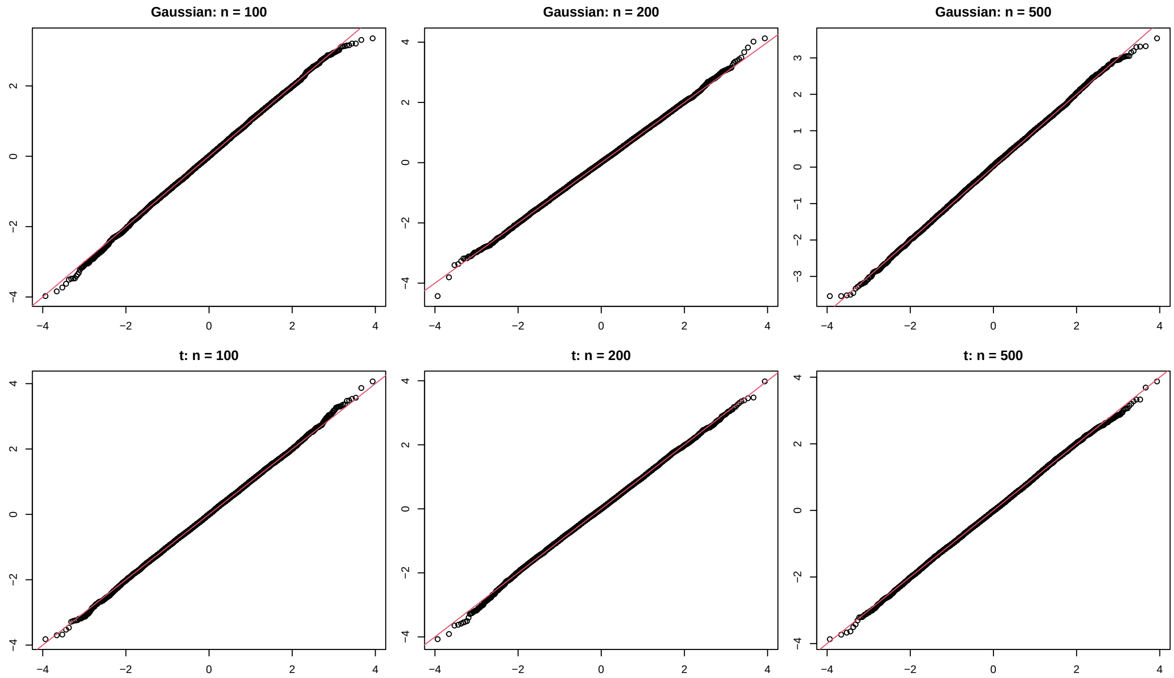


Figure C.16: (T2) Q-Q plot of the sample quantiles of $Z_{b,k,i}$, $i \in [p_k]$, $k \in [K]$, $b \in [B]$ ($y$-axis), against the quantiles from the standard normal distribution ($x$-axis) over varying $n \in \{100, 200, 500\}$ (left to right) and the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$, top to bottom) over $B = 100$ realisations per setting. In each plot, the $y = x$ line is given in red.

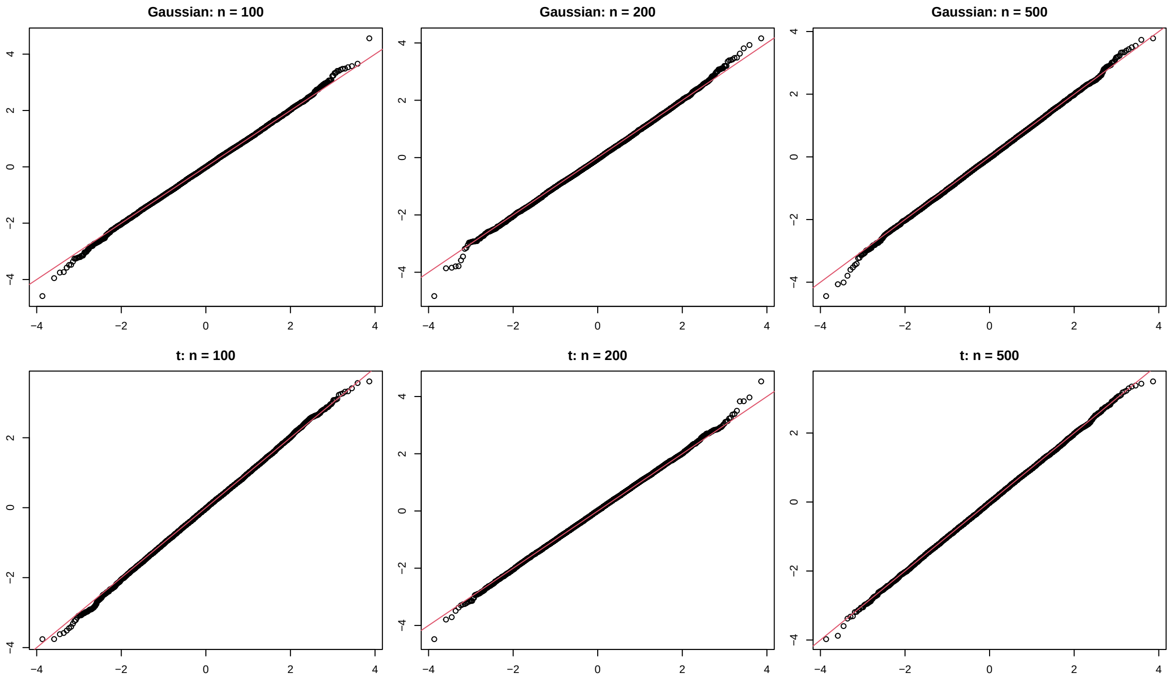


Figure C.17: (T3) Q-Q plot of the sample quantiles of $Z_{b,k,i}$, $i \in [p_k]$, $k \in [K]$, $b \in [B]$ ($y$-axis), against the quantiles from the standard normal distribution ($x$-axis) over varying $n \in \{100, 200, 500\}$ (left to right) and the distributions for $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ (Gaussian and $t_3$, top to bottom) over $B = 100$ realisations per setting. In each plot, the $y = x$ line is given in red.

### C.1.5 Additional simulation results

**Stronger serial dependence.** We investigate the performance of Trunc against its competitors when the degree of serial dependence in the tensor time series is stronger. For this, we focus on the model (T3) with fixed $n = 200$ and $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ generated from scaled $t_3$ distributions, where stronger serial dependence is introduced to $\mathcal{F}_t$ by considering $\phi \in \{0.7, 0.9\}$ in (C.1). See Figures C.18 and C.18 where we observe that the proposed Trunc exhibits good performance regardless of $\phi$ and thus is insensitive to the degree of serial dependence.

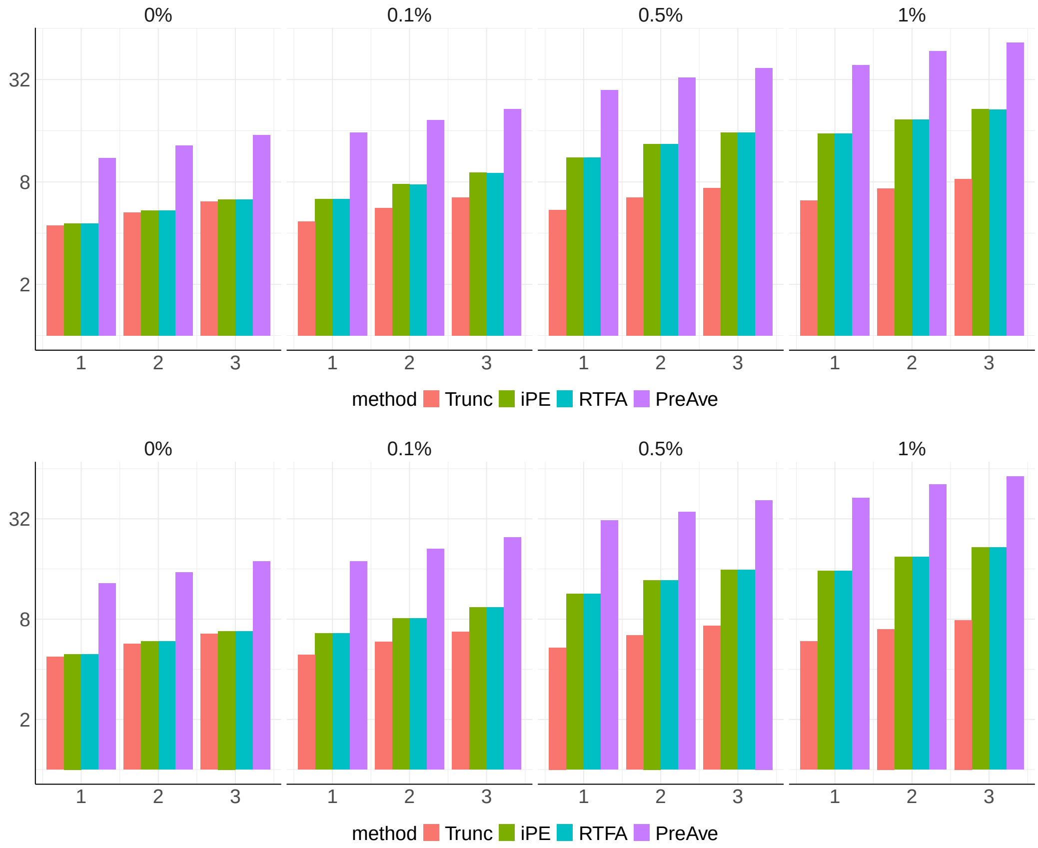


Figure C.18: (T3) with $(p_1, p_2, p_3) = (20, 30, 40)$: Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, iPE, RTFA and PreAve averaged over 100 realisations per setting, over varying $\varrho \in \{0, 0.1, 0.5, 1\}$ when $n = 200$ and $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ are generated from scaled $t_3$ distributions, with $\phi$ in (C.1) set to be $\phi = 0.7$ (top) and $\phi = 0.9$ (bottom). The $y$-axis is on a log scale and all errors have been scaled for better presentation.

**Convergence of the iteratively projected loading estimators.** We investigate the convergence behaviour of the iterative estimator proposed in Section 3.2. Focusing on the model (T3), we fix $n = 200$ and the distributions of $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ to be the (scaled) $t_3$, and report the loading estimation errors as measured in (C.2) for $\breve{\boldsymbol{\Lambda}}_k^{[0]} = \widehat{\boldsymbol{\Lambda}}_k$ (initial estimator) and $\breve{\boldsymbol{\Lambda}}_k^{[\iota]}$

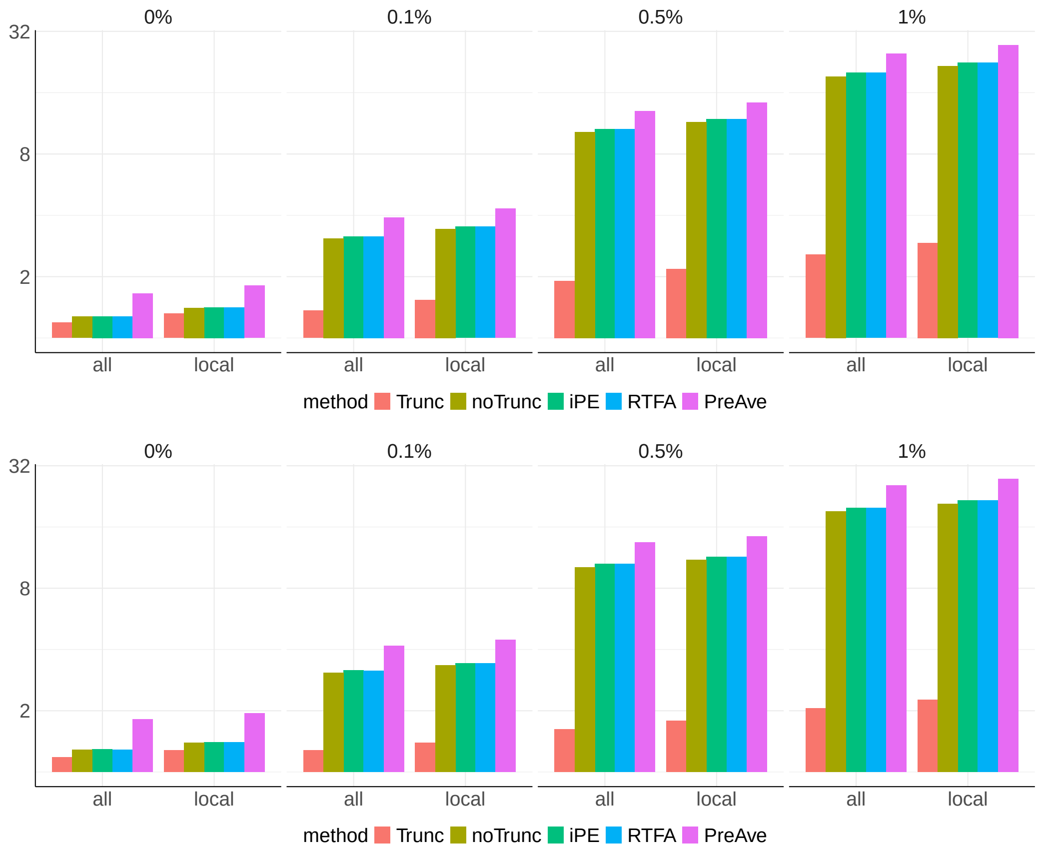


Figure C.19: (T3) with $(p_1, p_2, p_3) = (20, 30, 40)$: Common component estimation errors measured as in (C.3) for each mode ($x$-axis) for Trunc, noTrunc, iPE, RTFA and PreAve averaged over 100 realisations per setting, over varying $\varrho \in \{0, 0.1, 0.5, 1\}$ when $n = 200$ and $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ are generated from scaled $t_3$ distributions, with $\phi$ in (C.1) set to be $\phi = 0.7$ (top) and $\phi = 0.9$ (bottom). The $y$-axis is on a log scale and all errors have been scaled for better presentation.

for $\iota \in \{1, 2, \ldots, 10\}$ (we omit the dependence on $\tau$ for simplicity), see Figures C.20–C.21. The results indicate that after the initial large drop between the errors of $\breve{\boldsymbol{\Lambda}}_k^{[0]}$ and $\breve{\boldsymbol{\Lambda}}_k^{[1]}$, and the slight one between those of $\breve{\boldsymbol{\Lambda}}_k^{[1]}$ and $\breve{\boldsymbol{\Lambda}}_k^{[2]}$, there is little difference in the performance of $\breve{\boldsymbol{\Lambda}}_k^{[\iota]}$, $\iota \in \{2, \ldots, 10\}$, which supports our proposal of using the twice iterated estimator.

**$\boldsymbol{\Sigma}_k$ as Toeplitz matrix.** Recall the data generating process for $\boldsymbol{\xi}_t$ from Appendix C.1.1 (when no outlier is present):

$$\mathsf{vec}(\boldsymbol{\xi}_t) = \psi \cdot \mathsf{vec}(\boldsymbol{\xi}_{t-1}) + \sqrt{1 - \psi^2} \cdot \mathsf{vec}(\mathcal{V}_t),$$

where $\mathcal{V}_t \in \mathbb{R}^{p_1 \times p_2 \times p_3}$ such that $\mathsf{vec}(\mathcal{V}_t) = \otimes_{k=K}^{1} \boldsymbol{\Sigma}_k^{1/2} \mathbf{v}_t$ and the entries of $\mathbf{v}_t$ are i.i.d. To further understand the impact of cross-sectional correlations on our estimators, we con-

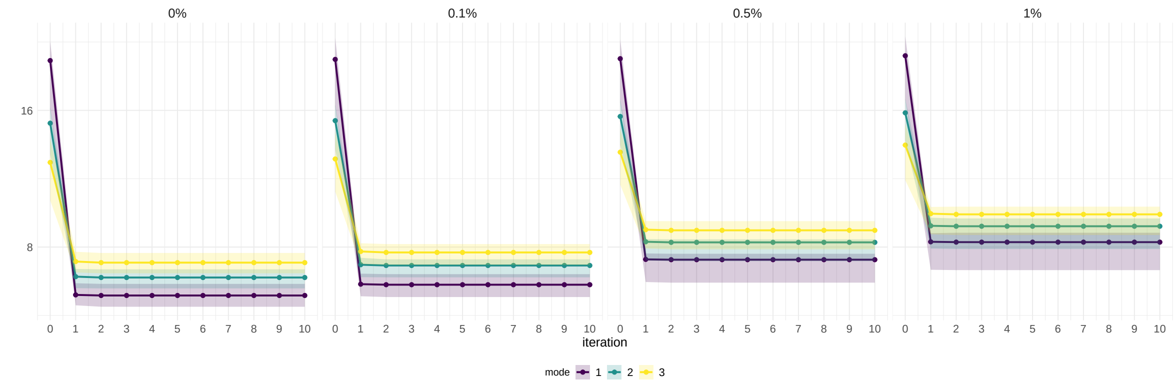


Figure C.20: (T3) with $(p_1, p_2, p_3) = (20, 30, 40)$, $n = 200$ and $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ generated from scaled $t_3$ distributions: Loading estimation errors measured as in (C.2) for $\breve{\mathbf{\Lambda}}_k^{[\iota]}$ for $\iota \in \{0, 1, \ldots, 10\}$ ($x$-axis) over 100 realisations per setting, with varying $\varrho \in \{0, 0.1, 0.5, 1\}$ (left to right). We display mean error curves with the shaded regions representing the interquartile range. The $y$-axis is on a log scale and all errors have been scaled for better presentation.

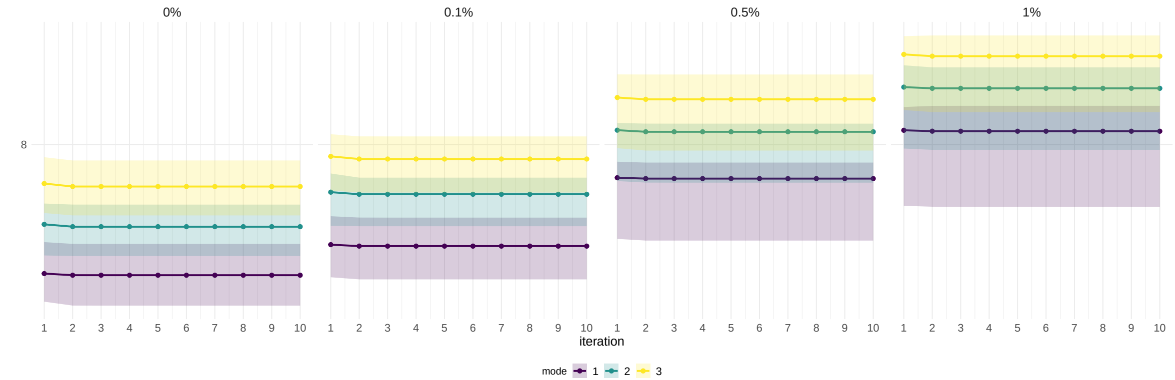


Figure C.21: Zoomed in version of Figure C.20 where we display the estimation errors from $\breve{\mathbf{\Lambda}}_k^{[\iota]}$ for $\iota \in \{1, \ldots, 10\}$ ($x$-axis).

sider the situations where $\mathbf{\Sigma}_k$ is a Toeplitz matrix, i.e. $\mathbf{\Sigma}_k = [\varphi^{|i-i'|}, i, i' \in [p_k]]$, with $\varphi \in \{0.5, 0.7, 0.9\}$; such a scenario better conforms to the strongly mixing random field condition in Assumption 5, compared to the one considered in the main text. We fix $n = 200$, $(p_1, p_2, p_3) = (20, 30, 40)$, $(r_1, r_2, r_3) = (3, 3, 3)$ and generate $\mathcal{F}_t$ and $\mathbf{v}_t$ from scaled $t_3$ distributions. In addition to 'Trunc', which refers to the twice-iterated estimator $\breve{\mathbf{\Lambda}}_k^{[2]}(\tau)$, we also present the results from the ten-times-iterated estimator $\breve{\mathbf{\Lambda}}_k^{[10]}(\tau)$ which is referred to as 'mTrunc', see Figure C.23 for the results.

We observe that in the presence of weaker cross-sectional correlations in $\boldsymbol{\xi}_t$ with $\varphi \in \{0.5, 0.7\}$, Trunc is performing as competitively as, or better than the competitors including mTrunc. This supports that in the presence of weak cross-sectional dependence, the additional iterations performed by mTrunc do not appear to make much difference, thus confirming the earlier observation on the fast convergence of the iteratively projected estimator. On the other hand, with $\varphi = 0.9$, we see a large difference between the outputs from Trunc and mTrunc. This is attributed to that such strong cross-sectional dependence in $\boldsymbol{\xi}_t$ may be regarded as being driven by extra 'factors'; indeed all factor number estimators including ours tend to over-estimate the number of factors with $\widehat{r}_k \geq 5$, when the true factor numbers are $(r_1, r_2, r_3) = (3, 3, 3)$, see also Figure C.22. In other words, the data generating process with $\varphi = 0.9$ may be considered as an extreme situation and Trunc shows invariably good performance in the presence of sufficiently weak cross-sectional correlations.

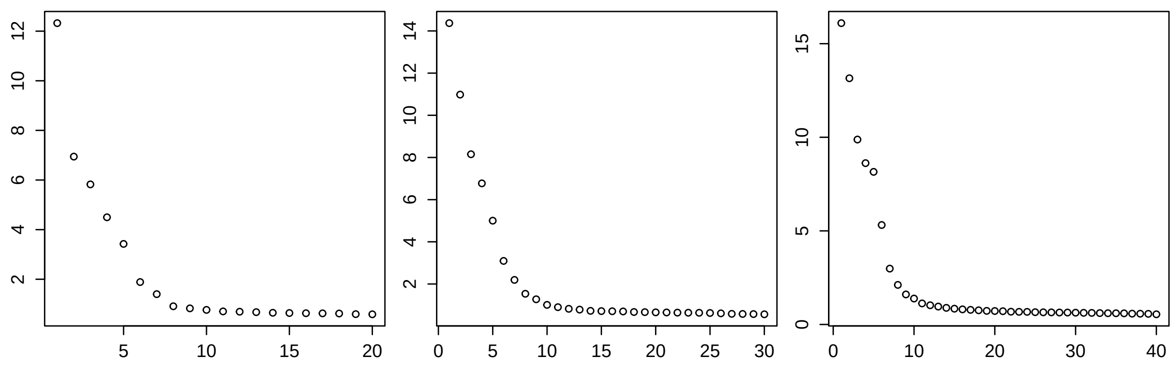


Figure C.22: Toeplitz scenarios with $n = 200$, $(p_1, p_2, p_3) = (20, 30, 40)$ and $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ are generated from scaled $t_3$ distributions: The ordered eigenvalues from $\widehat{\boldsymbol{\Gamma}}^{(k)}(\tau)$ (initial estimator) for $k \in \{1, 2, 3\}$ (left to right) on a single realisation when $\varphi = 0.9$ and $\varrho = 0.1$.

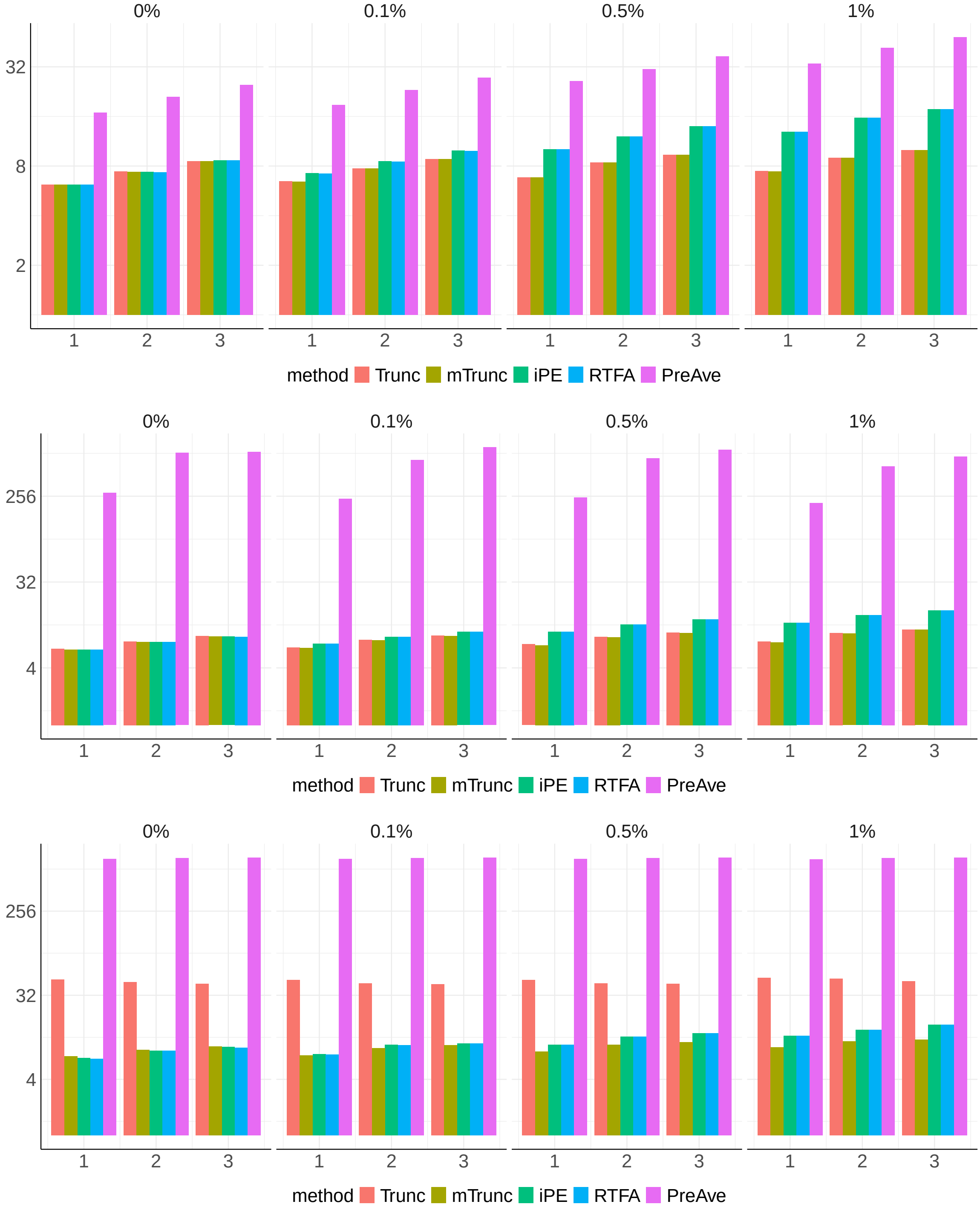


Figure C.23: Toeplitz scenarios with $n = 200$, $(p_1, p_2, p_3) = (20, 30, 40)$ and $\mathcal{F}_t$ and $\boldsymbol{\xi}_t$ are generated from scaled $t_3$ distributions: Loading estimation errors measured as in (C.2) for each mode ($x$-axis) for Trunc, mTrunc, iPE, RTFA and PreAve averaged over 100 realisations per setting, over varying $\varrho \in \{0, 0.1, 0.5, 1\}$ (left to right) and $\varphi \in \{0.5, 0.7, 0.9\}$ (top to bottom). The $y$-axis is on a log scale and all errors have been scaled for better presentation.

## C.2 Vector time series

### C.2.1 Set-up

**Data generation.** We adopt the following vector time series factor model ($K = 1$) from Ahn and Horenstein (2013):

$$X_{it}^\circ = \sum_{j=1}^{r} \lambda_{ij} f_{jt} + \xi_{it}, \quad \xi_{it} = \sqrt{\frac{1-\rho^2}{1+2J\beta^2}} e_{it},$$

$$e_{it} = \rho e_{i,t-1} + (1-\beta) v_{it} + \beta \cdot \sum_{\ell=\max(i-J,1)}^{\min(i+J,p)} v_{\ell t}, \quad i \in [p],\, t \in [n],$$

where $r = 3$ and $\lambda_{ij} \sim_{\text{iid}} \mathcal{N}(0,1)$. We set $\rho = \beta = J = 0$ in the 'independent' scenario while $(\rho, \beta, J) = (0.5, 0.2, \max(10, p/20))$ in the 'dependent' scenario. Following He et al. (2025), we consider the five models for the generation of $f_{jt}$ and $v_{it}$:

(V1) $f_{jt}$ and $v_{it}$ are i.i.d. random variables from the standard normal distribution.

(V2) $f_{jt}$ and $v_{it}$ are i.i.d. random variables from the scaled $t_3$ distribution such that $\mathsf{Var}(f_{jt}) = \mathsf{Var}(v_{it}) = 1$.

(V3) $f_{jt}$ are generated as in (V1) while $v_{it}$ are generated as in (V2).

(V4) $f_{jt}$ and $v_{it}$ are i.i.d. random variables from symmetric $\alpha$-stable distribution with the skewness parameter set at 0, the scale parameter at 1, the location parameter at 0 and the index parameter $\alpha = 1.9$ (we use the R package `stabledist` (Wuertz et al., 2016) for data generation).

(V5) $f_{jt}$ are i.i.d. random variables from skewed $t_3$ distribution with the slant parameter 20 (we use the R package `sn` (Azzalini, 2023) for data generation), while $v_{it}$ are generated as in (V4).

For the stable distribution considered in (V4) and (V5), the second moment does not exist with the choice of $\alpha = 1.9$.

With $X_{it}^\circ$ generated as above, we consider the situation where the observed data $X_{it}$ are contaminated by outliers. For this, we randomly select $\mathcal{O} \subset [p] \times [n]$ with its cardinality $|\mathcal{O}| = [\varrho np]$. Then for $(i,t) \in \mathcal{O}$, we set $X_{it} = s_{it} \cdot U_{it}$ with $s_{it} \sim_{\text{iid}} \{-1, 1\}$ and $U_{it} \sim_{\text{iid}} \text{Unif}[Q+12, Q+15]$ with $Q$ set to be the $\max(1 - 100/(np), 0.999)$-quantile of $|X_{it}^\circ|$, while $X_{it} = X_{it}^\circ$ otherwise; if $\varrho = 0$, we have $X_{it} = X_{it}^\circ$ for all $i$ and $t$. We vary $n, p \in \{100, 200, 500\}$ and $\varrho \in \{0, 0.1, 0.5, 1\} \times 10^{-2}$.

**Performance assessment.** To assess the factor loadings space estimation performance, for any estimator $\widehat{\mathbf{\Lambda}}$ we compute

$$\text{Err}_{\Lambda} = \sqrt{1 - \text{tr}\left(\Pi_{\widehat{\mathbf{\Lambda}}}\Pi_{\mathbf{\Lambda}}\right)/r}, \quad \text{where} \quad \Pi_{\mathbf{A}} = \mathbf{A}(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top. \tag{C.4}$$

To evaluate the quality in common component estimation, for any estimator $\widehat{\boldsymbol{\chi}}_t$ we evaluate

$$\text{Err}_{\chi}(\mathcal{T}) = \frac{\sum_{t\in\mathcal{T}} |\widehat{\boldsymbol{\chi}}_t - \boldsymbol{\chi}_t|_2^2}{\sum_{t\in\mathcal{T}} |\boldsymbol{\chi}_t|_2^2} \tag{C.5}$$

with $\mathcal{T} = [n]$ ('all') and $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local').

**Competitors.** In applying the proposed truncation-based estimator (hereafter referred to as 'Trunc'), we select the truncation parameters $\tau$ and $\kappa$ as described in Section 3.5, and obtain the estimators $\widehat{\mathbf{\Lambda}}(\tau) = \sqrt{p}\,\widehat{\mathbf{E}}(\tau)$ and $\widehat{\boldsymbol{\chi}}_t(\tau,\kappa) = \widehat{\mathbf{E}}(\tau)\widehat{\mathbf{E}}(\tau)^\top\mathbf{X}_t^{\mathbf{t}}(\kappa)$. For comparison, we consider an 'oracle' version of the proposed truncation-based methodology where we utilise the true (unobservable) loading matrix and common component for the selection of truncation parameters. Additionally, we include the classical PC-based estimator ('PCA') as well as the method proposed by He et al. (2022) for high-dimensional elliptical factor model ('RTS'), and the two methods based on minimising vector-variate ('HPCA') and element-wise ('IHR') Huber losses proposed by He et al. (2025), all implemented in the R package `HDRFA` (He et al., 2023).

### C.2.2 Results

For each setting, we generate 100 realisations and report the average and the standard deviation (in brackets) of the evaluation metrics in (C.4)–(C.5), see Figures C.24–C.29. We observe that different robust methods perform the best in different scenarios; IHR shows good performance across many scenarios, closely followed by Trunc. On the other hand, for common component estimation, we see overwhelming evidence favouring Trunc over other competitors in some scenarios such as (V4) and (V5) which supports the use of data truncation for factor estimation.

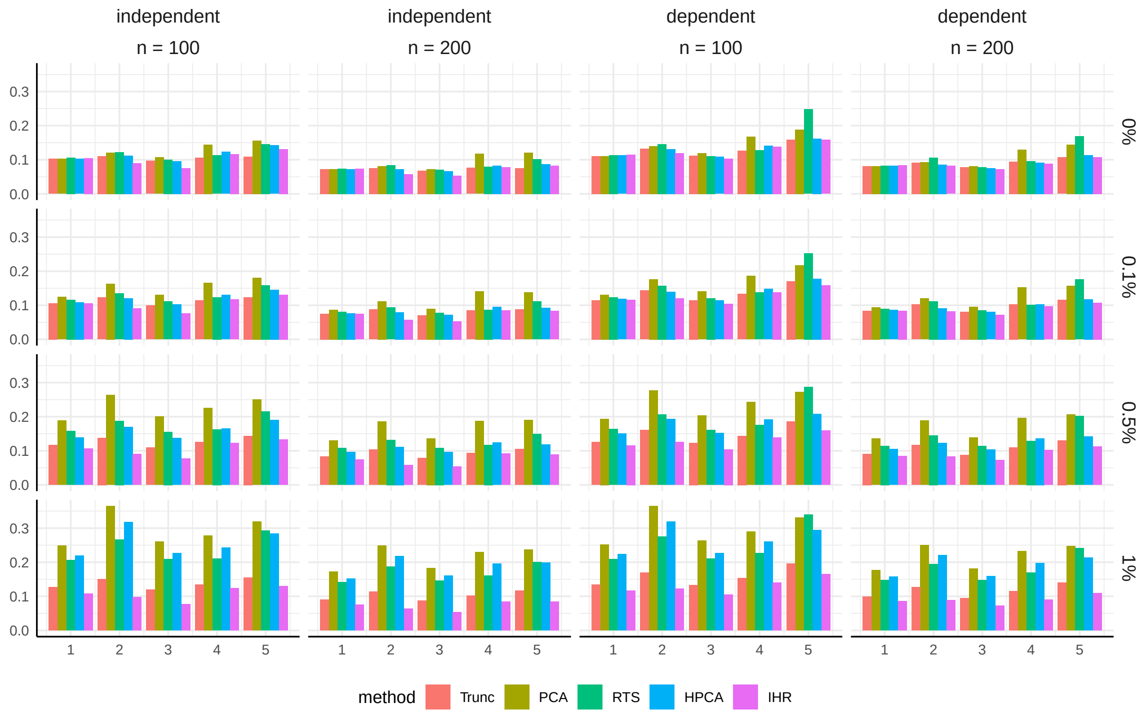


Figure C.24: Loading estimation errors measured as in (C.4) in different scenarios (V1)–(V5) ($x$-axis) for Trunc, PCA, RTS, HPCA and IHR over varying $n$ ($\{100, 200\}$), with and without temporal dependence in the idiosyncratic component and the percentage of outliers ($\{0, 0.1, 0.5, 1\}$), averaged over 100 realisations per setting. Here, $p = 100$.

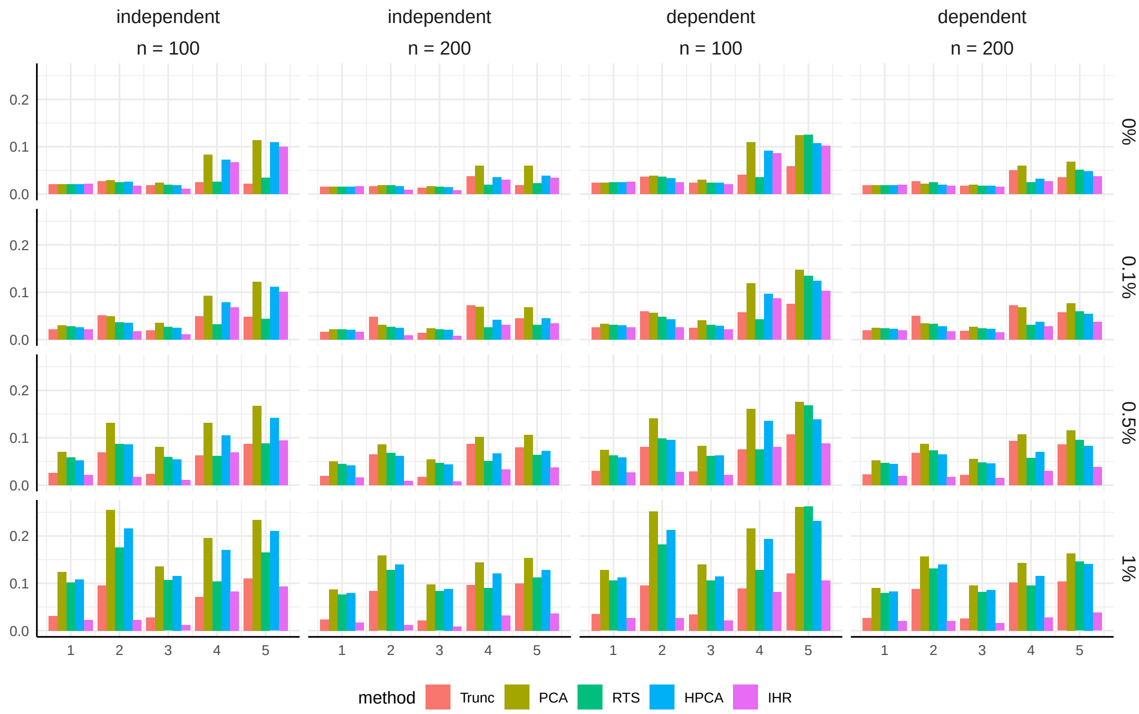


Figure C.25: Common component estimation errors measured as in (C.5) with $\mathcal{T} = [n]$ ('all') in different scenarios (V1)–(V5) ($x$-axis) for Trunc, PCA, RTS, HPCA and IHR over varying $n$ ($\{100, 200\}$), with and without temporal dependence in the idiosyncratic component and the percentage of outliers ($\{0, 0.1, 0.5, 1\}$), averaged over 100 realisations per setting. Here, $p = 100$.

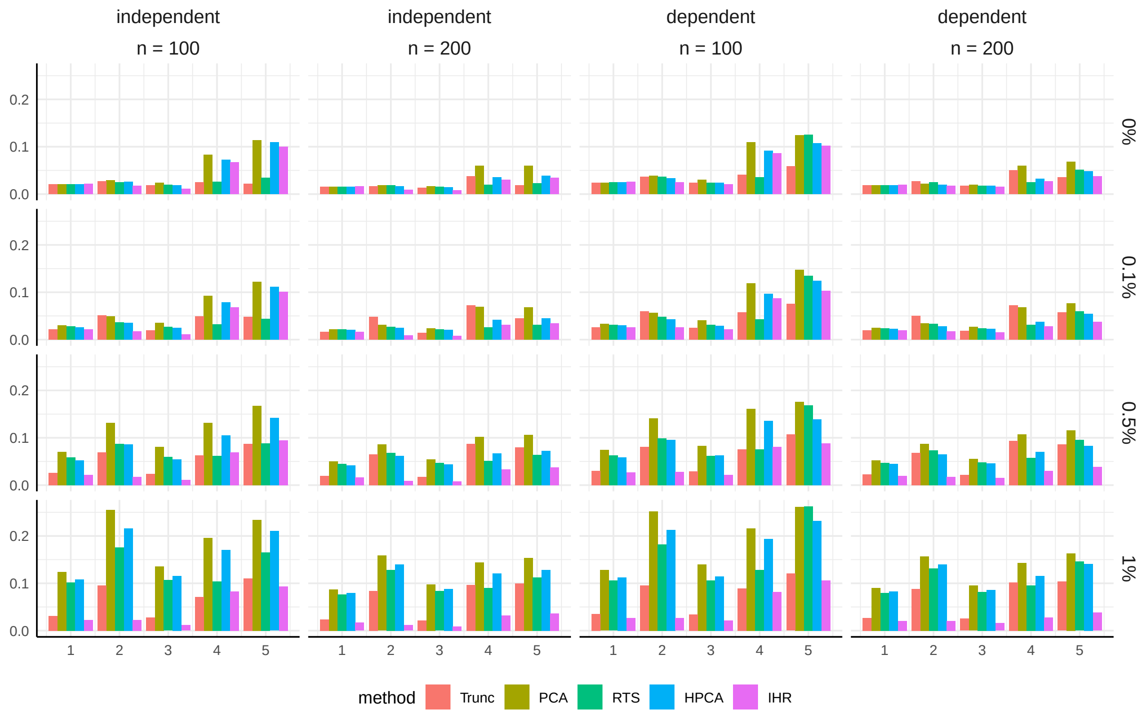


Figure C.26: Common component estimation errors measured as in (C.5) with $\mathcal{T} = \{n - 10 + 1, \ldots, n\}$ ('local') in different scenarios (V1)–(V5) ($x$-axis) for Trunc, PCA, RTS, HPCA and IHR over varying $n$ ($\{100, 200\}$), with and without temporal dependence in the idiosyncratic component and the percentage of outliers ($\{0, 0.1, 0.5, 1\}$), averaged over 100 realisations per setting. Here, $p = 100$.

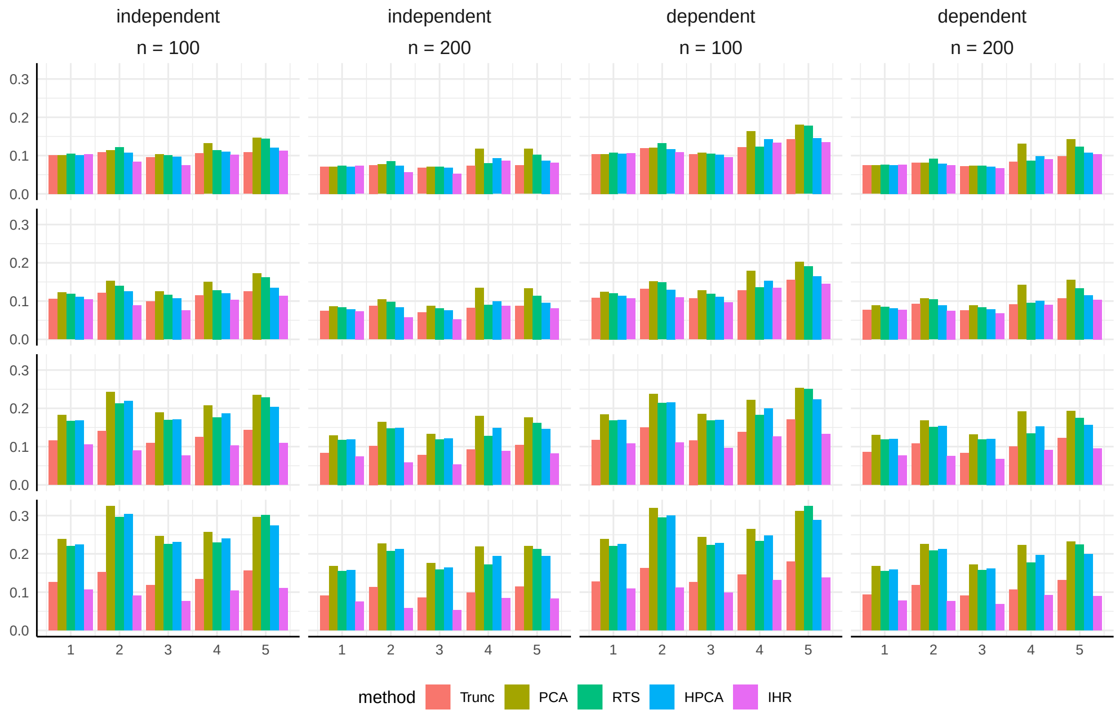


Figure C.27: Loading estimation errors measured as in (C.4) in different scenarios (V1)–(V5) ($x$-axis) for Trunc, PCA, RTS, HPCA and IHR over varying $n$ ($\{100, 200\}$), with and without temporal dependence in the idiosyncratic component and the percentage of outliers ($\{0, 0.1, 0.5, 1\}$), averaged over 100 realisations per setting. Here, $p = 200$.

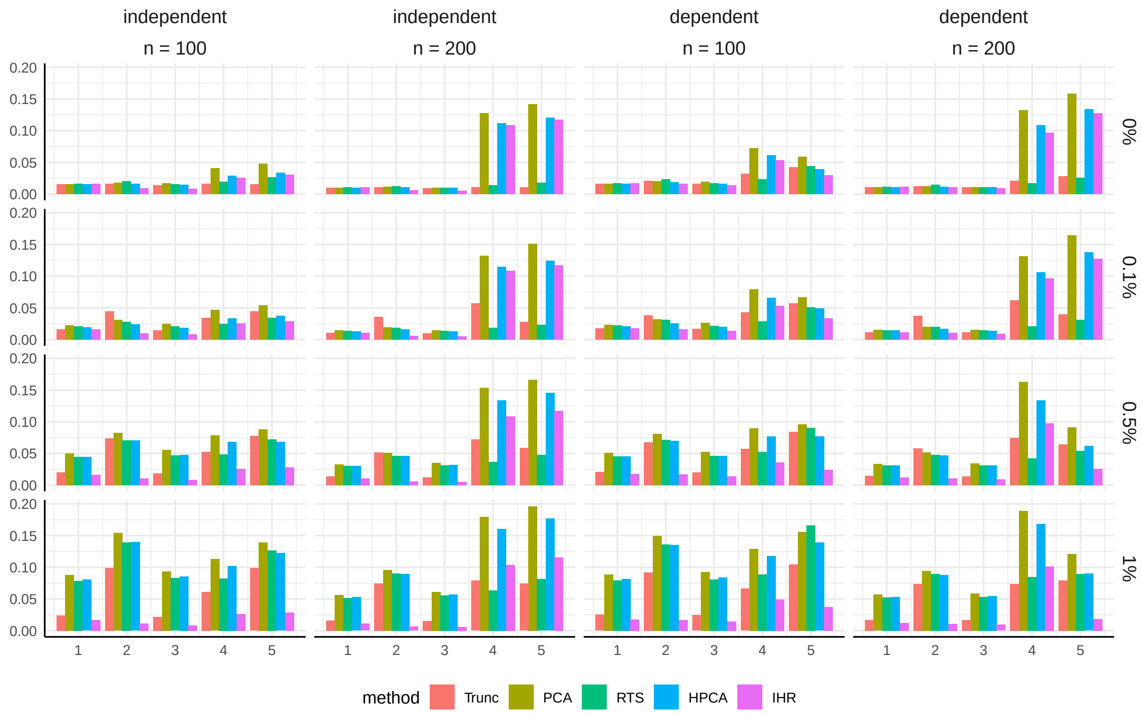


Figure C.28: Common component estimation errors measured as in (C.5) with $\mathcal{T} = [n]$ ('all') in different scenarios (V1)–(V5) ($x$-axis) for Trunc, PCA, RTS, HPCA and IHR over varying $n$ ($\{100, 200\}$), with and without temporal dependence in the idiosyncratic component and the percentage of outliers ($\{0, 0.1, 0.5, 1\}$), averaged over 100 realisations per setting. Here, $p = 200$.

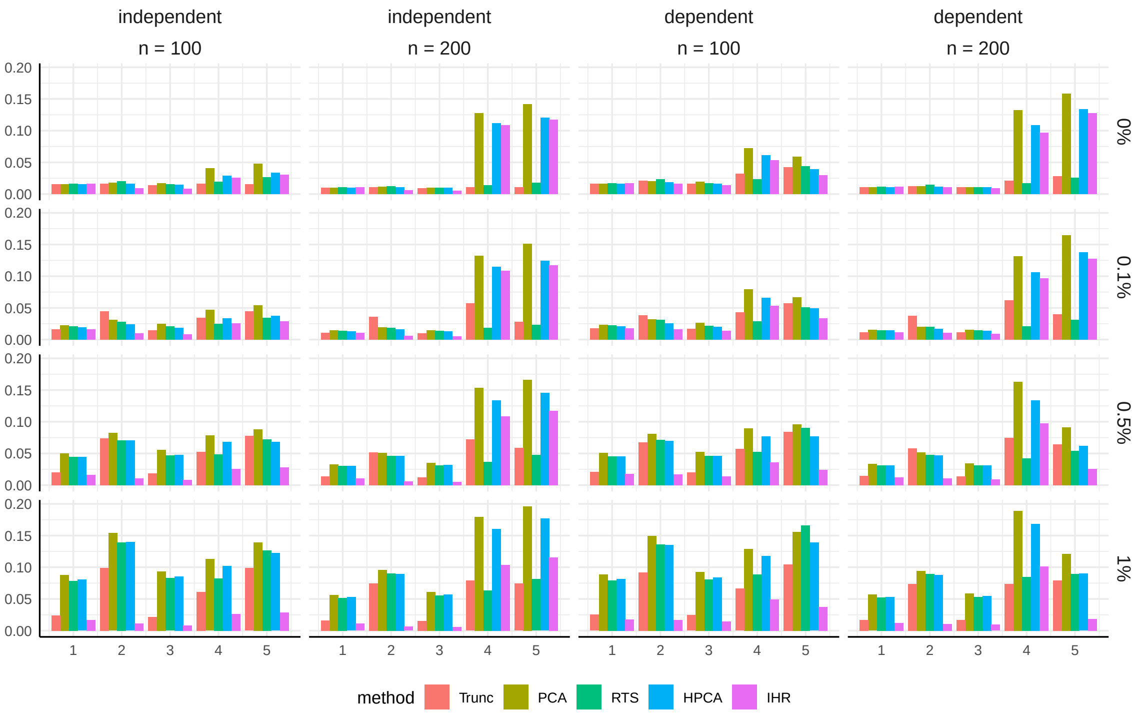


Figure C.29: Common component estimation errors measured as in (C.5) with $\mathcal{T} = \{n-10+1, \ldots, n\}$ ('local') in different scenarios (V1)–(V5) ($x$-axis) for Trunc, PCA, RTS, HPCA and IHR over varying $n$ ($\{100, 200\}$), with and without temporal dependence in the idiosyncratic component and the percentage of outliers ($\{0, 0.1, 0.5, 1\}$), averaged over 100 realisations per setting. Here, $p = 200$.

# D Additional empirical results

## D.1 Euro Area macroeconomic data

Continuing with EA-MD analysed in Section 6, Figure D.1 plots the output from the ratio-based factor number estimation as $\tau$ varies.

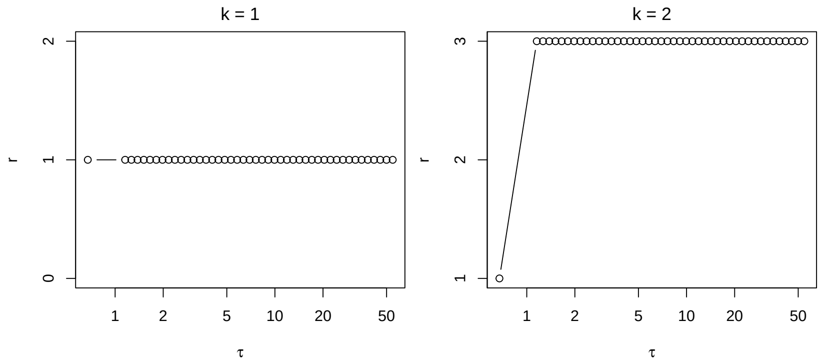


Figure D.1: EA-MD: Factor number estimators ($y$-axis) against the varying values of the truncation parameter $\tau$ ($x$-axis, in log scale) for $k = 1$ (left) and $k = 2$ (right).

Figure D.2 plots the output from the CV procedure described in Section 3.5 for the truncation parameter selection.

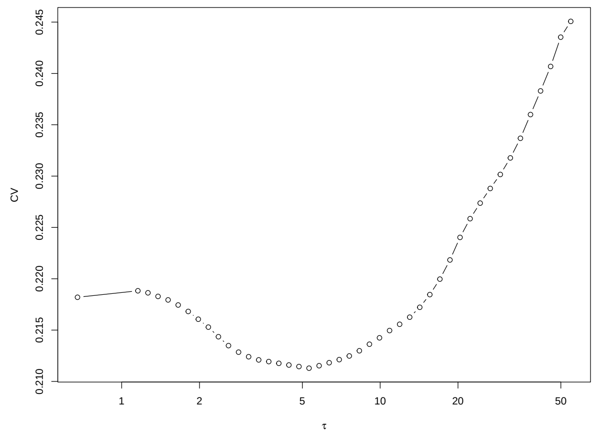


Figure D.2: EA-MD: CV measure in (9) with varying $\tau$ ($x$-axis, in log scale); $\tau_{\mathrm{CV}} \approx 5.306$.

Complementing Figure 4 that visualises the loading matrices obtained with and without truncation, Table D.1 summarises their ranges. Focusing on the estimators from Trunc, we observe that the elements of $\breve{\mathbf{\Lambda}}_1^{[2]}(\tau)$ are of the same sign across the 8 countries, indicating that the three factor time series aggregating the 37 indicators, are loaded onto the 8 countries in the same direction. This shows the evidence of a common EA factor, and that the correlations among the economic variables are similar within the 8 countries considered. Overall, Italy, Spain, Netherlands, and France have the larger loadings out of the 8 countries. The elements of $\breve{\mathbf{\Lambda}}_2^{[2]}(\tau)$ exhibit clusterings based on the grouping of the macroeconomic indicators.

Specifically, the first factor is strongly related to various confidence indicators such as the Economic Sentiment Indicator (ESENTIX) and, to a lesser extent, to the unemployment rate (UNETOT). The second factor is strongly related to Industrial Production (IDs starting with IP), and thus to real economic activity. The third factor is a nominal one, strongly related to price indexes such as the overall Harmonized Index of Consumer Prices (HICPOV).

Table D.1: EA-MD: The range of the elements of $\breve{\mathbf{\Lambda}}_k^{[2]}(\tau)$, $k \in \{1, 2\}$, and that from the Varimax rotated $\breve{\mathbf{\Lambda}}_2^{[2]}(\tau)$, for Trunc and noTrunc.

| | $\breve{\mathbf{\Lambda}}_1^{[2]}(\tau)$ | $\breve{\mathbf{\Lambda}}_2^{[2]}(\tau)$ | $\breve{\mathbf{\Lambda}}_2^{[2]}(\tau)$ after rotation |
|---|---|---|---|
| Trunc | $(0.622, 1.266)$ | $(-1.103, 2.645)$ | $(-1.158, 2.659)$ |
| noTrunc | $(0.432, 1.638)$ | $(-1.145, 2.945)$ | $(-1.172, 2.928)$ |

**Forecasting exercise.** We perform a forecast exercise to further illustrate the effect of truncation, whose result is summarised in Table 2. Denoting by $N = 236$ the number of observations prior to 2022-01 and $n = 257$ the total number of observations, we sequentially produce a one-step ahead forecast for the two indicators, the growth rate of the industrial production manufacturing index (IPMN) and the difference of the core consumer prices inflation (HICPNEF), on $t \in \{N+1, \ldots, n\}$, each time using all the preceding observations as the training data. For this, fixing $(\widehat{r}_1, \widehat{r}_2) = (1, 3)$, we estimate the loadings and the factors with and without the proposed truncation, fit a vector autoregressive (VAR) model to the vectorised factors with its order selected by Akaike information criterion, generate their one-step ahead forecast and then combine the latter with the loading estimates to produce the final forecast. With some abuse of notation, we denote the forecast at given $\mathbf{i} \in [p_1] \times \{\text{IPMN, HICPNEF}\}$ and time $t$ by $\widehat{X}^{\star}_{\mathbf{i},t|t-1}$, where $\star \in \{\text{Trunc, noTrunc}\}$ denotes whether the truncation is performed or not throughout the estimation procedure. Then, we inspect the forecast error

$$\text{Err}_{\mathbf{i},t}(\star) = \left| \widehat{X}^{\star}_{\mathbf{i},t|t-1} - X_{\mathbf{i},t} \right|, \quad \star \in \{\text{Trunc, noTrunc}\}. \tag{D.1}$$

As each new observation arrives, we update the loading and the factor estimates and the VAR fitted to the estimated factors, which is repeated until the end of the dataset is reached.

## D.2 US macroeconomic data

FRED-MD is a large, monthly frequency, macroeconomic database maintained by the Federal Reserve Bank of St. Louis (McCracken and Ng, 2016) and has been analysed frequently in the time series factor modelling literature as a benchmark. We use the dataset spanning the period from 1960 to 2023 ($n = 767$). Removing the time series with missing observations, we have $p = 111$ variables in total. All time series are individually transformed to stationarity as suggested in McCracken and Ng (2016), see their Appendix I for full details of variable-specific transformations. Further, to handle the heterogeneity in the scale of the variables, we center and standardise each time series. We opt to use the mean and the standard deviation in place of the more robust statistics such as the median and the mean absolute deviation (MAD), as some indicators have MAD very close to zero over some rolling windows considered in the forecasting exercise described below.

We consider two approaches for factor number estimation: The first one, proposed in Alessi et al. (2010), applies the information criteria of Bai and Ng (2002) to the subsets of the data of varying dimensions and sample sizes in order to mitigate the arbitrariness in the choice of constants applied to the penalty. The second is the ratio-based estimator in Section 3.4 with $\bar{r} = \min(\lfloor p/2 \rfloor, 20) = 20$. Since their performance depends on the choice of the truncation parameter $\tau$, we examine the outputs from the two methods with varying values of $\tau$, see Figure D.3. For the first approach, the three information criteria of Bai and Ng (2002) with different penalties attain consensus over the most $\tau$ values at $\widehat{r} = 5$, while the second one favours $\widehat{r} = 3$. For comparison, the Huber PCA-based estimator (He et al., 2025) gives $\widehat{r} = 4$. Below we present the forecasting results with $\widehat{r} \in \{3, 5\}$ and demonstrate that regardless of its choice, the proposed estimator performs competitively.

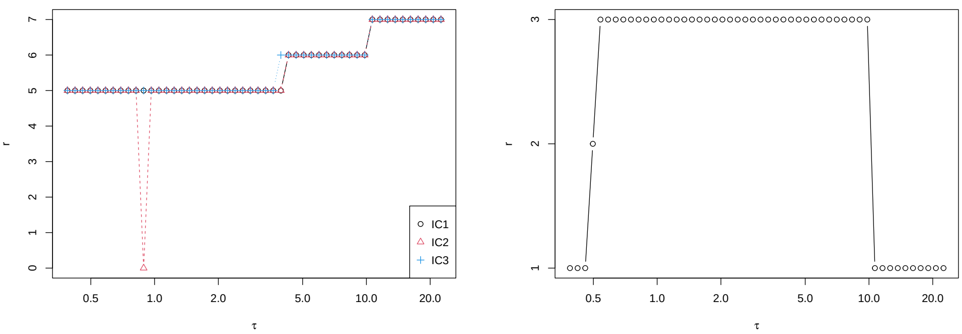


Figure D.3: FRED-MD: Factor number estimators ($y$-axis) against the varying values of the truncation parameter $\tau$ ($x$-axis, in log scale) generated by the information criterion-based method of Alessi et al. (2010) (left) and by the ratio-based estimator in Section 3.4 (right).

**Forecasting exercise.** We compare the forecasting performance of the proposed truncation-based estimator against the non-robust counterpart. We also considered the method proposed by He et al. (2025) but do not report the results due to its numerical convergence issues.

Let $X_{i,t}$ denote the variable of interest to forecast, which may be any one of the $p$ variables. For given window size $T$ and forecasting horizon $h \geq 1$, let $\widehat{X}_{i,t+h|T}$ denote the $h$-step ahead forecast of the $i$-th variable based on the past observations at $\{t - T + 1, \ldots, t\}$, obtained as an estimator of the best linear predictor $\mathrm{Proj}(X_{i,t+h} | \mathcal{X}_u,\ u \leq t)$, where $\mathrm{Proj}(\cdot|\mathbf{z})$ denotes the linear projection operator onto the space spanned by $\mathbf{z}$ (Stock and Watson, 2002). In line with the tail-robust method described in Section 3, we propose the estimator

$$\widehat{X}_{i,t+h|T}(\tau,\kappa) = \boldsymbol{\varphi}_i^\top (\widehat{\boldsymbol{\Gamma}}_t(\tau,h))^\top \widehat{\mathbf{E}}_t(\tau)(\widehat{\mathbf{M}}_t(\tau))^{-1}\widehat{\mathcal{F}}_t(\kappa),$$

where $\widehat{\boldsymbol{\Gamma}}_t(\tau,h) = T^{-1}\sum_{u=t-T+1}^{t-h} \mathcal{X}_u^{\mathsf{t}}(\tau)(\mathcal{X}_{u+h}^{\mathsf{t}}(\tau))^\top$, $\widehat{\mathbf{E}}_t(\tau)$ and $\widehat{\mathbf{M}}_t(\tau)$ contain the leading $\widehat{r}$ eigenvectors and eigenvalues of $\widehat{\boldsymbol{\Gamma}}_t(\tau,0)$, respectively, and $\boldsymbol{\varphi}_i$ is the unit vector with its $i$-th element set to one; see also Barigozzi et al. (2024). We refer to such an estimator combined with the truncation parameters chosen via CV (see Section 3.5) as 'Trunc', while the one combined with $\tau = \kappa = \infty$ (i.e. no truncation) by 'noTrunc'. Following Trucíos et al. (2021), we compare their forecasting performance in a rolling window-based exercise with $T = 12 \times 10$ (10 years).

At given $t$, we aggregate the forecasting error over the forecasting horizons $h \in [24]$, as

$$\mathrm{Err}_{it}(\star) = \frac{1}{24}\sum_{h=1}^{24}\left|\widehat{X}^{\star}_{i,t+h|T} - X_{i,t+h}\right|, \quad \star \in \{\mathrm{Trunc}, \mathrm{noTrunc}\},$$

and also write $\overline{\mathrm{Err}}_i(\star) = (n-T-24)^{-1}\sum_{t=T+1}^{n-24}\mathrm{Err}_{it}(\star)$. Out of the $p = 111$ variables, we have $\overline{\mathrm{Err}}_i(\mathrm{Trunc}) < \overline{\mathrm{Err}}_i(\mathrm{noTrunc})$ for 106 variables with $\widehat{r} = 3$, and the inequality holds for all $p$ variables with $\widehat{r} = 5$, see Figure D.4.

To further investigate whether the two methods perform significantly differently at any point during the forecasting exercise, we employ the fluctuation test proposed by Giacomini and Rossi (2010) (implemented in the R package `murphydiagram` (Jordan and Krueger, 2019)), and focus on two variables, the growth rate of the industrial production total index (INDPRO) and the difference of the consumer prices inflation index (all items, CPIAUCSL). The test requires a tuning parameter $\mu \in (0,1)$ which determines the bandwidth $\lceil \mu(n-24-T) \rceil$ over which the moving average of the forecasting error differences is produced, and it is recommended that $\mu \geq 0.2$; we consider $\mu \in \{0.2, 0.3, 0.4\}$. Figures D.5–D.8 display the difference in the forecasting errors and the results from the fluctuation test with $\widehat{r} \in \{3,5\}$ and $\mu \in \{0.2, 0.3, 0.4\}$. Regardless of the choice of $\widehat{r}$ or $\mu$, we observe that the two approaches perform significantly differently in forecasting over certain periods where Trunc is favoured over noTrunc. Such intervals correspond to where the estimation involves the data from the Great Financial Crisis (2007–09) and Covid-19 pandemic (2020–21). The raw difference in forecasting error shows that entering the crisis period, $\mathrm{Err}_{it}(\mathrm{Trunc}) - \mathrm{Err}_{it}(\mathrm{noTrunc})$ may initially increase as the forecasting target itself is highly anomalous, but it is followed by a sharp decrease indicating that the robust approach benefits from truncating such anomalous observations in estimation.

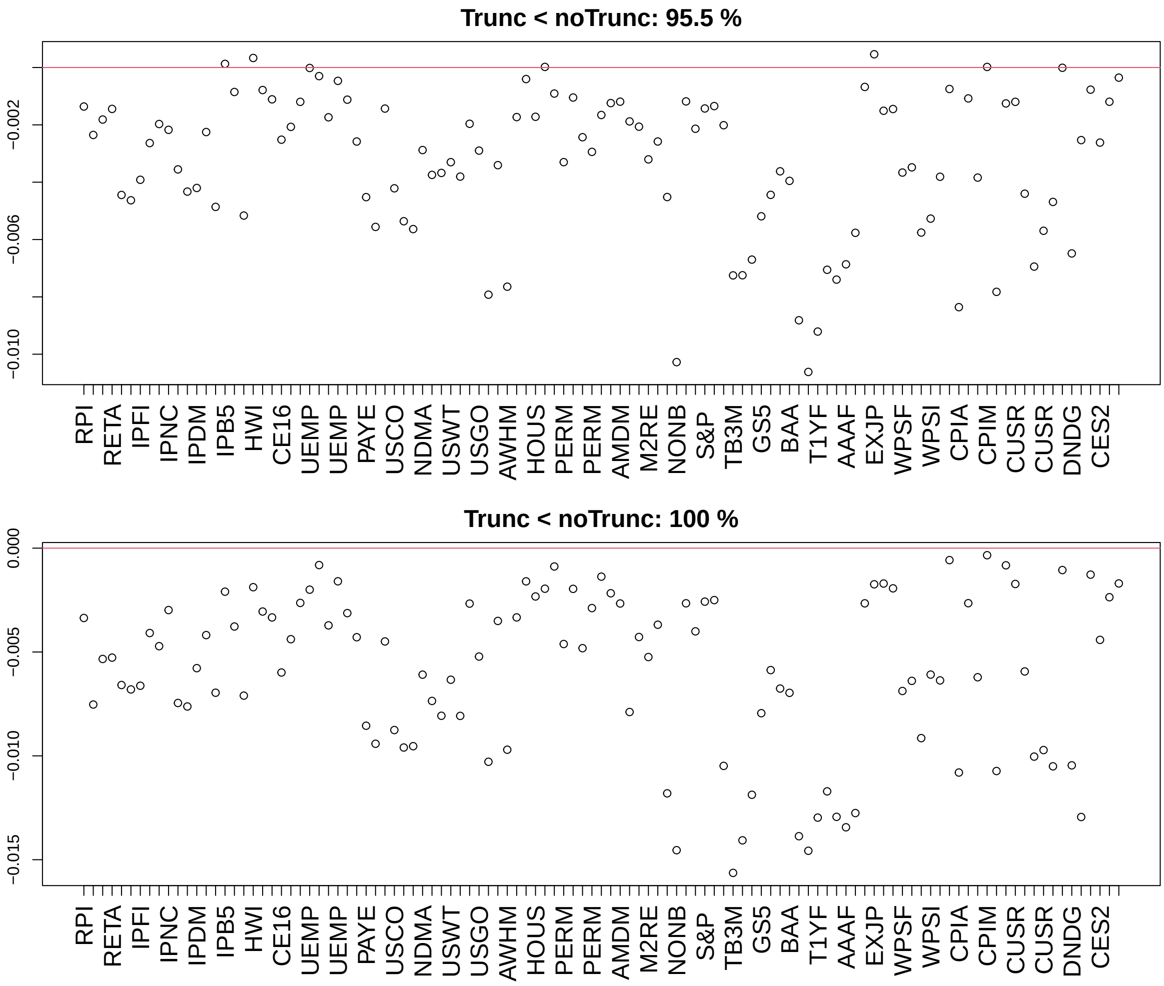


Figure D.4: $\overline{\mathrm{Err}}_i(\mathrm{Trunc}) - \overline{\mathrm{Err}}_i(\mathrm{noTrunc})$ for all $i \in [p]$ ($x$-axis) after standardisation by the overall standard deviation, with $\widehat{r} = 3$ (top) and $\widehat{r} = 5$ (bottom). The first four letters for each variable's ticker are given as the $x$-axis labels. Horizontal lines are drawn at $y = 0$.

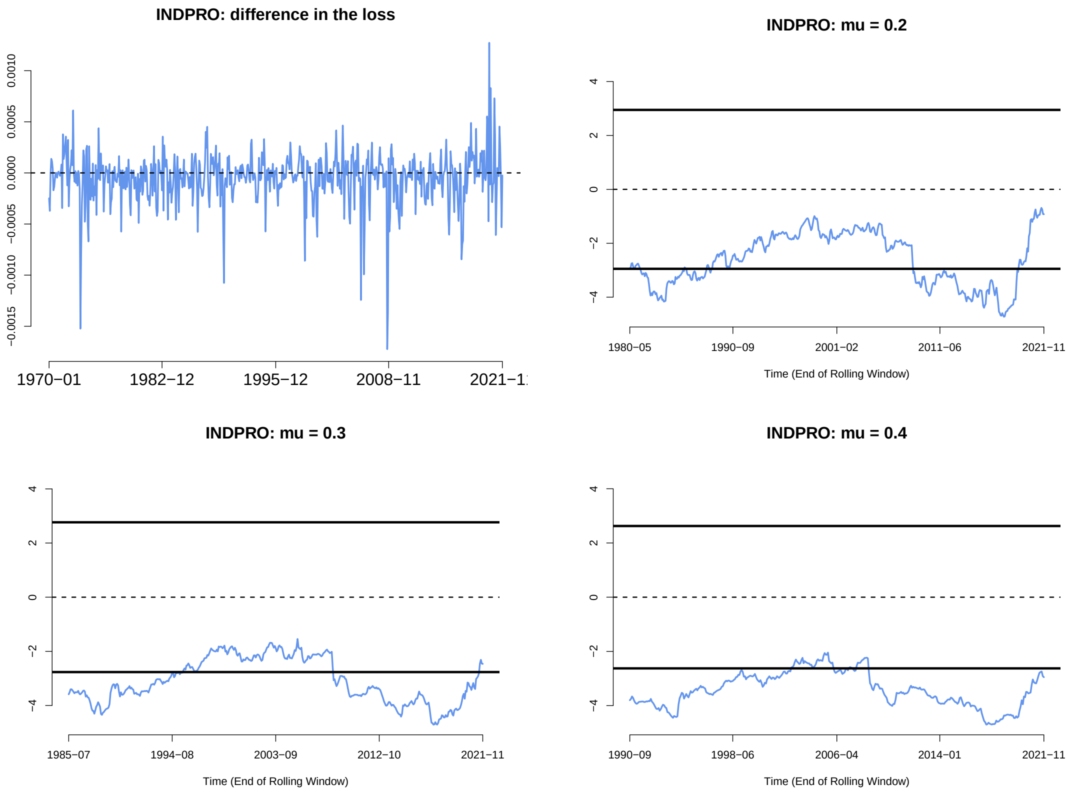


Figure D.5: Industrial production total index: $\text{Err}_{it}(\text{Trunc}) - \text{Err}_{it}(\text{noTrunc})$, $T+1 \leq t \leq n-24$ (top left) and the corresponding fluctuation test statistics computed with $\mu \in \{0.2, 0.3, 0.4\}$ along with the two-sided critical values at the significance level $\alpha = 0.1$. When the fluctuation test statistic falls below the lower solid line, Trunc outperforms the noTrunc and vice versa. Here, we set $\widehat{r} = 3$.

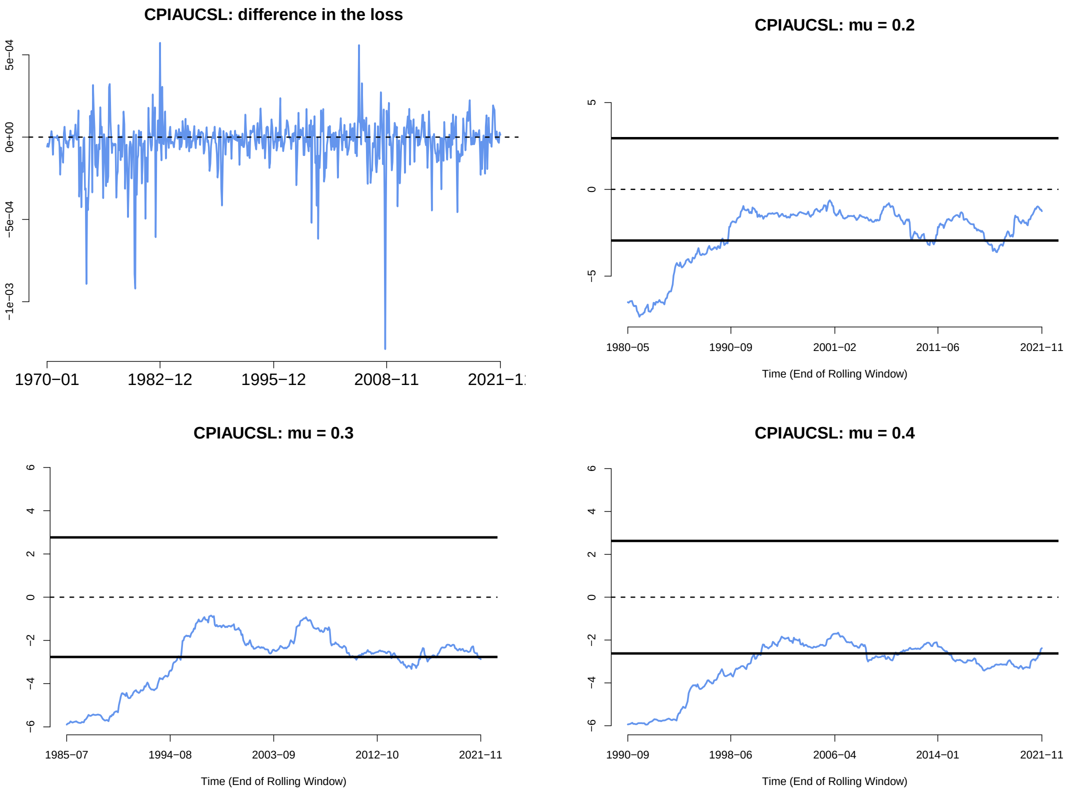


Figure D.6: Consumer prices index (all items): $\text{Err}_{it}(\text{Trunc}) - \text{Err}_{it}(\text{noTrunc})$, $T+1 \leq t \leq n-24$ (top left) and the corresponding fluctuation test statistics computed with $\mu \in \{0.2, 0.3, 0.4\}$ along with the two-sided critical values at the significance level $\alpha = 0.1$. When the fluctuation test statistic falls below the lower solid line, Trunc outperforms the noTrunc and vice versa. Here, we set $\widehat{r} = 3$.

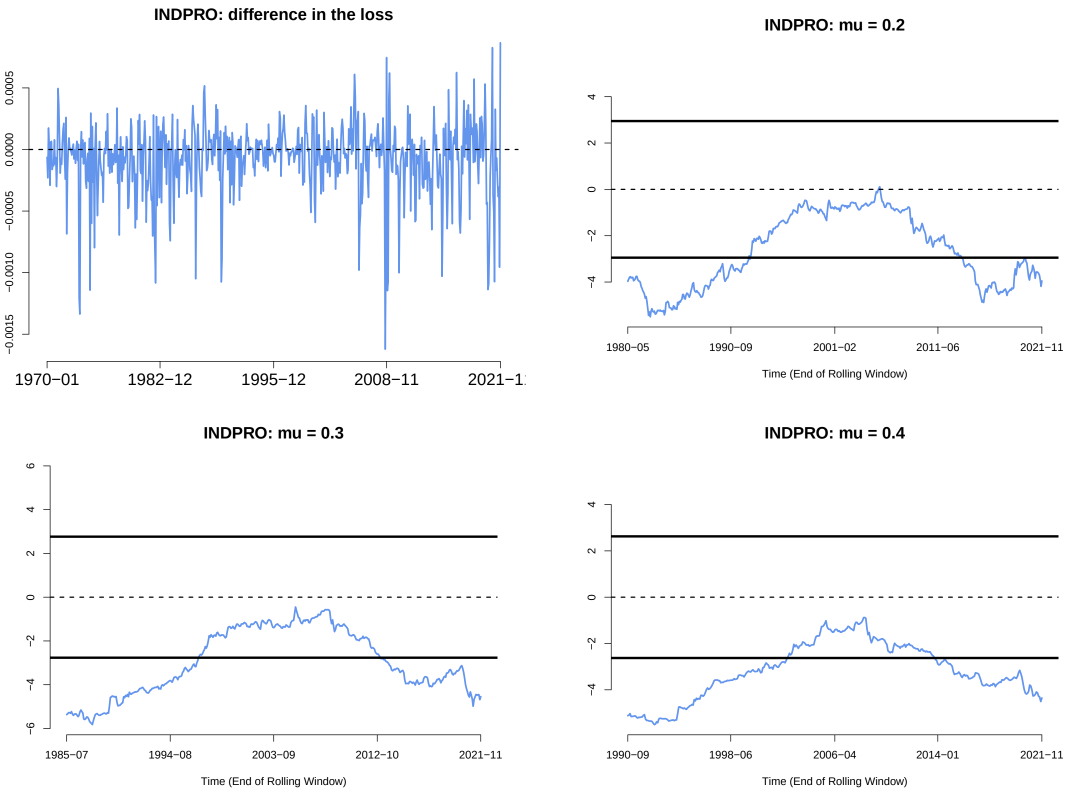


Figure D.7: Industrial production total index: $\text{Err}_{it}(\text{Trunc}) - \text{Err}_{it}(\text{noTrunc})$, $T+1 \leq t \leq n-24$ (top left) and the corresponding fluctuation test statistics computed with $\mu \in \{0.2, 0.3, 0.4\}$ along with the two-sided critical values at the significance level $\alpha = 0.1$. When the fluctuation test statistic falls below the lower solid line, Trunc outperforms the noTrunc and vice versa. Here, we set $\widehat{r} = 5$.

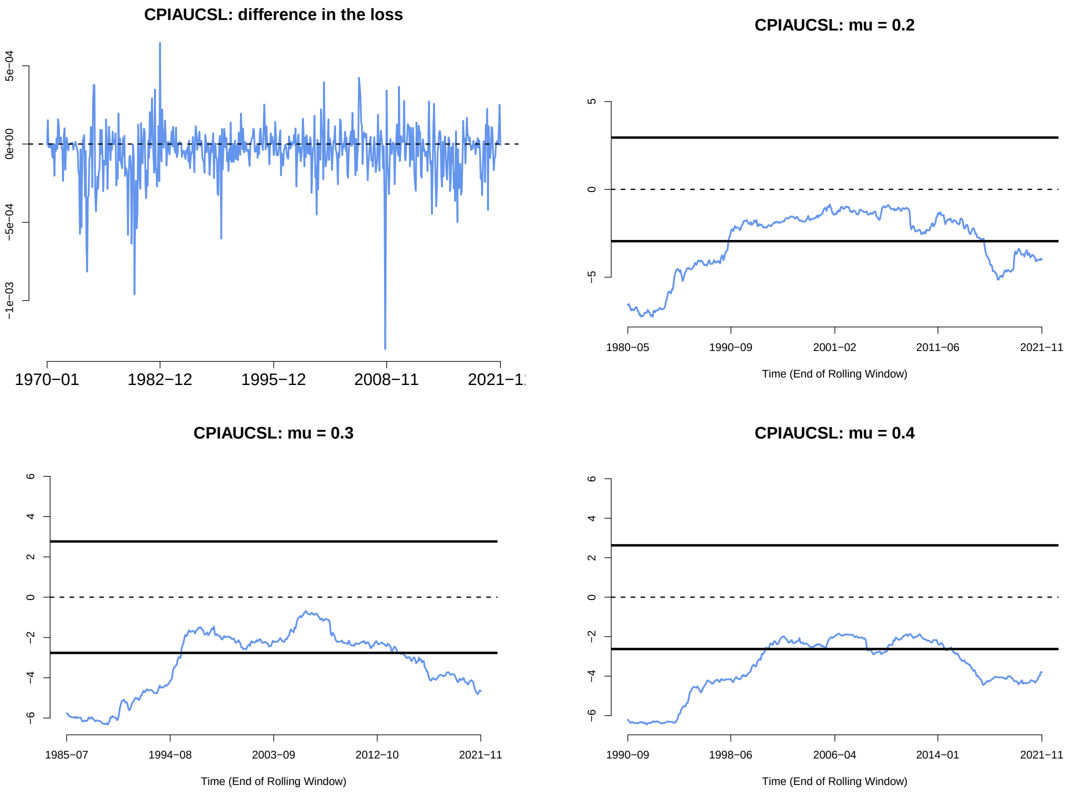


Figure D.8: Consumer prices index (all items): $\mathrm{Err}_{it}(\mathrm{Trunc}) - \mathrm{Err}_{it}(\mathrm{noTrunc})$, $T+1 \leq t \leq n-24$ (top left) and the corresponding fluctuation test statistics computed with $\mu \in \{0.2, 0.3, 0.4\}$ along with the two-sided critical values at the significance level $\alpha = 0.1$. When the fluctuation test statistic falls below the lower solid line, Trunc outperforms the noTrunc and vice versa. Here, we set $\widehat{r} = 5$.

## References


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

Alessi, L., Barigozzi, M., and Capasso, M. (2010). Improved penalization for determining the number of factors in approximate static factor models. *Statistics & Probability Letters*, 80:1806–1813.

Azzalini, A. A. (2023). *The R package* `sn`*: The skew-normal and related distributions such as the skew-t and the SUN*. Università degli Studi di Padova, Italia. R package version 2.1.1.

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

Barigozzi, M., Cho, H., and Owens, D. (2024). Fnets: Factor-adjusted network estimation and forecasting for high-dimensional time series. *Journal of Business & Economic Statistics*, 42(3):890–902.

Barigozzi, M., He, Y., Li, L., and Trapani, L. (2022). Statistical inference for large-dimensional tensor factor models by iterative projections. *arXiv preprint arXiv:2206.09800*.

Barigozzi, M., He, Y., Li, L., and Trapani, L. (2023). Robust tensor factor analysis. *arXiv preprint arXiv:2303.18163*.

Billingsley, P. (1995). *Probability and Measure*. John Wiley & Sons.

Giacomini, R. and Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, 25(4):595–620.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application*. Academic Press.

He, Y., Kong, X., Yu, L., and Zhang, X. (2022). Large-dimensional factor analysis without moment constraints. *Journal of Business & Economic Statistics*, 40(1):302–312.

He, Y., Li, L., Liu, D., and Zhou, W. (2023). *HDRFA: High-Dimensional Robust Factor Analysis*. R package version 0.1.4.

He, Y., Li, L., Liu, D., and Zhou, W.-X. (2025). Huber principal component analysis for large-dimensional factor models. *Journal of Econometrics*, 249:105993.

Jordan, A. and Krueger, F. (2019). *murphydiagram: Murphy Diagrams for Forecast Comparisons*. R package version 0.12.2.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Merlevede, F. and Peligrad, M. (2020). Functional clt for nonstationary strongly mixing processes. *Statistics & Probability Letters*, 156:108581.

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

Trucíos, C., Mazzeu, J. H., Hotta, L. K., Pereira, P. L. V., and Hallin, M. (2021). Robustness and the general dynamic factor model with infinite-dimensional space: identification, estimation, and forecasting. *International Journal of Forecasting*, 37(4):1520–1534.

Wang, D. and Tsay, R. S. (2023). Rate-optimal robust estimation of high-dimensional vector autoregressive models. *The Annals of Statistics*, 51(2):846–877.

Wuertz, D., Maechler, M., and Rmetrics core team members (2016). *stabledist: Stable Distribution Functions*. R package version 0.7-1.

Yu, Y., Wang, T., and Samworth, R. J. (2015). A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika*, 102:315–323.