Supplementary Material for
“Shrinkage Estimation of Network Spillovers with Factor Structured Errors”

Ayden Higgins
University of Surrey, UK

Federico Martellosio
University of Surrey, UK

December 4, 2020

Contents

B Identification

Further to the discussion at the start of Section 3, a few additional remarks are provided regarding identification. The identification condition usually encountered when establishing the consistency of extremum estimators requires that the limit of the objective function is uniquely maximised at the true parameter values. This is referred to as *extremum* identification, in order to distinguish it from the usual concept of identification, which requires that the distribution of the data be unique at the true population parameters. These two identification concepts are related, however, one does not generally imply the other. In a likelihood setting, extremum identification can be shown to be both necessary and sufficient for identification since $\mathbb{E}[\mathcal{L}(\boldsymbol{\theta})] < \mathbb{E}[\mathcal{L}(\boldsymbol{\theta}^0)]$ if and only if $\mathcal{L}(\boldsymbol{\theta}) < \mathcal{L}(\boldsymbol{\theta}^0)$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}^0 \in \boldsymbol{\Theta}$, when $\mathcal{L}(\boldsymbol{\theta})$ is a true likelihood. In the case of a quasi-likelihood, extremum identification can be shown to still imply identification, for a class of distributions with certain properties that includes the normal distribution; see Lee and Yu (2016).[1] Yet with a penalised quasi-likelihood, this implication is no longer guaranteed. Typically, however, the properties of penalty functions ensure that their effect diminishes asymptotically, in which case the limiting penalised objective function will reduce to its unpenalised counterpart, and establishing identification using the latter is then equivalent to doing so with the former. As mentioned in the main text, it is relatively easier to establish consistency directly in the present context, however, Assumption ID below sets out conditions that, in addition to Assumptions 1 and 2, can be used to formulate an explicit identification argument.

Assumption ID.

ID.1 $R \geq R^0$.

ID.2 $\mathbb{E}[\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}})\boldsymbol{\mathcal{Z}}]$ *is positive definite for all* $\boldsymbol{\Lambda} \in \mathbb{R}^{n\times R}$.

ID.3 $\mu_1(\mathbb{E}[\boldsymbol{\mathcal{Z}}'\boldsymbol{\mathcal{Z}}]) < \infty$.

Notice that these conditions are similar to Assumption 4.2 in the main text, and indeed, they are counterparts to those assumptions when considering explicitly the limiting objective function and thus share analogous intuition.

Proposition ID (Identification)**.** *Under Assumptions 1,2 and ID, the parameters* $\boldsymbol{\theta}^0, \sigma_0^2$ *and the product* $\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}$ *are identified.*

The proof of this proposition can be found in Appendix H. Under Assumptions 1, 2 and ID, the unpenalised expected log-likelihood is uniquely maximised at $\boldsymbol{\theta}^0, \sigma_0^2, \boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}$ which implies that distribution of the data must be unique at the true population parameters. Moreover, the true number of factors R^0 can be recovered from the rank of the matrix $\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}$. This establishes identification in the usual sense. With addition of Assumption 3, the limiting penalised objective function asymptotically reduces to its unpenalised counterpart, and hence the penalised objective function is also uniquely maximised at the true parameter values. One may ask if, for the present model, it is possible to follow the Bramoullé et al.

[1] Incidentally, the reverse is no longer true; identification no longer always implies extremum identification.

(2009) approach and derive identification conditions in terms of the network structure (see also Kwok, 2019, for the case of multiple weights matrices). Unfortunately, the answer seems to be negative, because the error factor structure makes it impossible to derive a reduced form free of fixed effects as in Bramoullé et al. (2009).

C Assumption 4.2 and Equations (13) and (14)

C.1 Assumption 4.2

Assumption 4.2 can be related to analogous conditions appearing elsewhere in the literature by means of the relationship

$$\begin{aligned}\inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R},\boldsymbol{F}^0\in\mathbb{R}^{T\times R^0}} \mu_P\left(\boldsymbol{\mathcal{H}}_1\right) &= \min_{\boldsymbol{\alpha}\in\mathbb{R}^P:||\boldsymbol{\alpha}||_2=1} \sum_{r=R+R^0+1}^{n} \mu_r\left(\frac{1}{nT}(\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{Z}})(\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{Z}})'\right) \\ &\leq \inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}} \mu_P\left(\boldsymbol{\mathcal{H}}_1\right), \end{aligned} \tag{C.1}$$

where $\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{Z}} \coloneqq \sum_{p=1}^{P}\alpha_p\boldsymbol{\mathcal{Z}}_p$. To establish (C.1), note that

$$\begin{aligned} &\inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R},\boldsymbol{F}^0\in\mathbb{R}^{T\times R^0}} \mu_P\left(\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}})\boldsymbol{\mathcal{Z}}\right) \\ &= \min_{\boldsymbol{\alpha}\in\mathbb{R}^P:||\boldsymbol{\alpha}||_2=1} \inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R},\boldsymbol{F}^0\in\mathbb{R}^{T\times R^0}} \frac{1}{nT}(\boldsymbol{\mathcal{Z}}\boldsymbol{\alpha})'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}})\boldsymbol{\mathcal{Z}}\boldsymbol{\alpha} \\ &= \min_{\boldsymbol{\alpha}\in\mathbb{R}^P:||\boldsymbol{\alpha}||_2=1} \inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R},\boldsymbol{F}^0\in\mathbb{R}^{T\times R^0}} \frac{1}{nT}\mathrm{tr}\left((\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{Z}})'\boldsymbol{M}_{\boldsymbol{\Lambda}}(\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{Z}})\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ &= \min_{\boldsymbol{\alpha}\in\mathbb{R}^P:||\boldsymbol{\alpha}||_2=1} \sum_{r=R+R^0+1}^{n} \mu_r\left(\frac{1}{nT}(\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{Z}})(\boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{Z}})'\right), \end{aligned} \tag{C.2}$$

where the last line follows from Lemma A.1 in Moon and Weidner (2017). The inequality in (C.1) follows since the minimum eigenvalue of $\inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}})\boldsymbol{\mathcal{Z}}$ can be no less than if one could also minimise over the space of true factors. The first line of (C.1) shows that Assumption 4.2 is equivalent to Assumption NC in Moon and Weidner (2015), which avoids mention of the unobservable population factors $\boldsymbol{F}^0$. Assumption A in Bai (2009) is equivalent to the requirement that $\inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}}\mu_P\left(\boldsymbol{\mathcal{H}}_1\right) > 0$ in the limit. It is clear from (C.1) that this is a weaker requirement than Assumption 4.2; the need for a stronger condition arises since the consistency result in the present paper assumes that the number of factors is not understated, rather than known, as in Bai (2009).

C.2 Equations (13) and (14)

Here, equations (13) and (14) are derived. The intuition is that Assumption 4.2 provides upper and lower bounds on variation in the data from which the inequalities (13) and (14) can be derived. It would usually be assumed that the matrix $\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{\mathcal{Z}}$ is positive definite in the limit. However, it is shown next that the eigenvalues of $\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{\mathcal{Z}}$ can be no less than $\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}})\boldsymbol{\mathcal{Z}}$ whereby Assumption 4.2 implies (13) and (14). Let $\boldsymbol{M} := \boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}}$. Since the Kronecker product of two symmetric and idempotent matrices is also symmetric and idempotent, both $\boldsymbol{M}$ and $\boldsymbol{P} := \boldsymbol{I}_{n\times T} - \boldsymbol{M}$ are symmetric and idempotent.[2] From Weyl's inequality, for two $n\times n$ symmetric matrices $\boldsymbol{A},\boldsymbol{B}$ of the same size $\mu_i(\boldsymbol{A})+\mu_n(\boldsymbol{B}) \leq \mu_i(\boldsymbol{A}+\boldsymbol{B})$ for $i = 1,..,n$ (e.g., Horn and Johnson, 2012, Corollary 4.3.15). As $\boldsymbol{\mathcal{Z}}'\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{Z}}'\boldsymbol{M}\boldsymbol{\mathcal{Z}}+\boldsymbol{\mathcal{Z}}'\boldsymbol{P}\boldsymbol{\mathcal{Z}}$, and all three of these matrices are real and symmetric, then for $p = 1, ..., P$,

$$\mu_p\left(\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{M}\boldsymbol{\mathcal{Z}}\right) + \mu_P\left(\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{P}\boldsymbol{\mathcal{Z}}\right) \leq \mu_p\left(\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{\mathcal{Z}}\right). \tag{C.3}$$

By Assumption 4.2, $\mu_P\left(\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{M}\boldsymbol{\mathcal{Z}}\right) > 0$ in the limit. Also $\mu_P\left(\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{P}\boldsymbol{\mathcal{Z}}\right) \geq 0$ since $\boldsymbol{P}$ is idempotent, and therefore $\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{P}\boldsymbol{\mathcal{Z}}$ must be positive semidefinite. Hence (14) follows from (C.3). Similarly, (13) follows from (C.3) since $\mu_1(\frac{1}{nT}\boldsymbol{\mathcal{Z}}'\boldsymbol{M}\boldsymbol{\mathcal{Z}}) < c$ in the limit.

D Proofs of Propositions 1 and 4

This Appendix provides a more detailed proof of Proposition 1 and a proof of Proposition 4.

Proof of Proposition 1. Due to the diverging number of parameters, this consistency proof follows the approach taken by Fan and Peng (2004). Let $\boldsymbol{u}$ be a $P\times 1$ vector, and $\mathcal{T}_{nT}(\boldsymbol{\theta}^0) := \{\boldsymbol{\theta}^0 + a_{nT}\boldsymbol{u} : ||\boldsymbol{u}||_2 \leq d\}$ be a closed ball centred at $\boldsymbol{\theta}^0$ with radius $a_{nT}d$. The objective is to show that for any $\epsilon > 0$ and sufficiently large n, T, there exists a large enough d such that

$$\Pr\left(\sup_{||\boldsymbol{u}||_2=d} \mathcal{Q}(\boldsymbol{\theta}^0 + a_{nT}\boldsymbol{u}) < \mathcal{Q}(\boldsymbol{\theta}^0)\right) \geq 1-\epsilon. \tag{D.1}$$

Because $\mathcal{T}_{nT}(\boldsymbol{\theta}^0)$ is compact, (D.1) implies that, as $n, T\to\infty$, there exists a local maximiser in the interior of $\mathcal{T}_{nT}(\boldsymbol{\theta}^0)$ with probability approaching 1, call this $\hat{\boldsymbol{\theta}}_L$, such that $||\hat{\boldsymbol{\theta}}_L-\boldsymbol{\theta}^0||_2 <$

[2]If $A = A'$ and $B = B'$, then $(A\otimes B)' = (A'\otimes B') = (A\otimes B)$; if $A = AA$, and $B = BB$ then $(A\otimes B)(A\otimes B) = A\otimes B$.

$a_{nT}||\boldsymbol{u}||_2 = O_P(a_{nT})$. First, however, the existence of an a_{nT}-consistent local maximiser of the unpenalised objective function, denoted $\tilde{\boldsymbol{\theta}}_L$, is demonstrated. This follows from showing

$$\Pr\left(\sup_{||\boldsymbol{u}||_2=d} \mathcal{L}(\boldsymbol{\theta}^0 + a_{nT}\boldsymbol{u}) < \mathcal{L}(\boldsymbol{\theta}^0)\right) \geq 1-\epsilon, \tag{D.2}$$

where the average concentrated quasi-likelihood is given by

$$\mathcal{L}(\boldsymbol{\theta}) \coloneqq \sup_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}} \left\{\frac{1}{n}\log(\det(\boldsymbol{S}(\boldsymbol{\rho}))) - \frac{1}{2}\log\left(\hat{\sigma}^2(\boldsymbol{\theta},\boldsymbol{\Lambda})\right)\right\}. \tag{D.3}$$

To begin, a lower bound for $\mathcal{L}(\boldsymbol{\theta}^0)$ is established. Evaluating (D.3) at $\boldsymbol{\theta}^0$ and substituting in the true data generating process yields

$$\mathcal{L}(\boldsymbol{\theta}^0) = \frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\log\left(\inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)\right\}\right). \tag{D.4}$$

Now,

$$\begin{aligned} 0 &\leq \inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)\right\} \\ &\leq \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) \\ &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t - \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}_t. \end{aligned} \tag{D.5}$$

By Assumption 1.1, $\mathbb{E}[\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t] = \sigma_0^2$ and thus, by the law of large numbers, $\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t = \sigma_0^2 + O_P\left(\frac{1}{\sqrt{nT}}\right)$. For the second term in (D.5),

$$\left|\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}_t\right| = \frac{1}{nT}|\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})| \leq \frac{1}{nT}R^0||\boldsymbol{\varepsilon}||_2^2 = O_P\left(\frac{1}{\min\{n,T\}}\right). \tag{D.6}$$

This gives the result that

$$\underline{\mathcal{L}}(\boldsymbol{\theta}^0) := \frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\log\left(\sigma_0^2 + O_P\left(\frac{1}{\min\{n,T\}}\right)\right) \leq \mathcal{L}(\boldsymbol{\theta}^0). \tag{D.7}$$

Consider $\sup_{||\boldsymbol{u}||_2=d}\{\mathcal{L}(\boldsymbol{\theta}^0 + a_{nT}\boldsymbol{u})\}$. Let $\ddot{\boldsymbol{u}} \coloneqq \arg\sup_{||\boldsymbol{u}||_2=d}\{\mathcal{L}(\boldsymbol{\theta}^0 + a_{nT}\boldsymbol{u})\}$. Partition $\ddot{\boldsymbol{u}}$ into two vectors, $\ddot{\boldsymbol{u}}_\rho$ and $\ddot{\boldsymbol{u}}_\beta$ with the former being $Q\times 1$ and the latter being $K\times 1$. Define $\ddot{\boldsymbol{\theta}} \coloneqq \boldsymbol{\theta}^0 + a_{nT}\ddot{\boldsymbol{u}}$, $\ddot{\boldsymbol{\rho}} \coloneqq \boldsymbol{\rho}^0 + a_{nT}\ddot{\boldsymbol{u}}_\rho$ and $\ddot{\boldsymbol{\beta}} \coloneqq \boldsymbol{\beta}^0 + a_{nT}\ddot{\boldsymbol{u}}_\beta$. One then has

$$\mathcal{L}(\ddot{\boldsymbol{\theta}}) = \frac{1}{n}\log(\det(\boldsymbol{S}(\ddot{\boldsymbol{\rho}}))) - \frac{1}{2}\log\left(\inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}}\hat{\sigma}^2(\ddot{\boldsymbol{\theta}},\boldsymbol{\Lambda})\right). \tag{D.8}$$

Next an upper bound for $\mathcal{L}(\ddot{\boldsymbol{\theta}})$ is derived. Substituting the true data generating process into (D.8) yields

$$\begin{aligned}\mathcal{L}(\ddot{\boldsymbol{\theta}}) &= \frac{1}{n}\log(\det(\boldsymbol{S}(\ddot{\boldsymbol{\rho}}))) - \frac{1}{2}\log\left(\inf_{\boldsymbol{\Lambda}\in\mathbb{R}^{n\times R}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}(\boldsymbol{X}_t\boldsymbol{\beta}^0 + \boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) - \boldsymbol{X}_t\ddot{\boldsymbol{\beta}})'\right.\right.\\ &\quad\left.\left.\times \boldsymbol{M}_{\boldsymbol{\Lambda}}(\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}(\boldsymbol{X}_t\boldsymbol{\beta}^0 + \boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) - \boldsymbol{X}_t\ddot{\boldsymbol{\beta}})\right\}\right)\\ &\leq \frac{1}{n}\log(\det(\boldsymbol{S}(\ddot{\boldsymbol{\rho}}))) - \frac{1}{2}\log\left(\inf_{\dot{\boldsymbol{\Lambda}}\in\mathbb{R}^{n\times(R+R^0)}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{X}_t\boldsymbol{\beta}^0 + \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t - \boldsymbol{X}_t\ddot{\boldsymbol{\beta}})'\right.\right.\\ &\quad\left.\left.\times \boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}(\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{X}_t\boldsymbol{\beta}^0 + \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t - \boldsymbol{X}_t\ddot{\boldsymbol{\beta}})\right\}\right), \qquad \text{(D.9)}\end{aligned}$$

where the last expression is obtained by also minimising with respect to $\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0$ and $\boldsymbol{F}^0$ since the value of the objective function can be no less than if one was also able to minimise over $\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0$ and $\boldsymbol{F}^0$. Lemma A.1 in Moon and Weidner (2017) then demonstrates the equivalence between this and the second expression as it appears in (D.9), where the expression is now minimised over $\dot{\boldsymbol{\Lambda}} \in \mathbb{R}^{n\times(R+R^0)}$ because the rank of $\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{F}^{0'} - \boldsymbol{\Lambda}\boldsymbol{F}'$ can be no greater than $R + R^0$. Applying Lemma A.2(i) to (D.9) gives

$$\begin{aligned}\mathcal{L}(\ddot{\boldsymbol{\theta}}) \leq \frac{1}{n}\log(\det(\boldsymbol{S}(\ddot{\boldsymbol{\rho}}))) - \frac{1}{2}\log&\left(\inf_{\dot{\boldsymbol{\Lambda}}\in\mathbb{R}^{n\times(R+R^0)}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}}) + \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)'\right.\right.\\ &\left.\left.\times\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}}) + \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)\right\}\right). \qquad \text{(D.10)}\end{aligned}$$

Expanding the term inside of the log in (D.10),

$$\begin{aligned}&\inf_{\dot{\boldsymbol{\Lambda}}\in\mathbb{R}^{n\times(R+R^0)}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}}) + \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}}) + \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)\right\}\\ &\geq \inf_{\dot{\boldsymbol{\Lambda}}\in\mathbb{R}^{n\times(R+R^0)}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}}))'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}})\right\} + \frac{2}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t\\ &\quad - \sup_{\dot{\boldsymbol{\Lambda}}\in\mathbb{R}^{n\times(R+R^0)}}\left\{\frac{2}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \ddot{\boldsymbol{\theta}})'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t\right\} + \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t\\ &\quad - \sup_{\dot{\boldsymbol{\Lambda}}\in\mathbb{R}^{n\times(R+R^0)}}\left\{\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t\right\}\\ &=: k_1 + ... + k_5. \qquad \text{(D.11)}\end{aligned}$$

Consider the probability order of terms $k_1, ..., k_5$.

$$
\begin{aligned}
k_1 &= a_{nT}^2 \inf_{\dot{\boldsymbol{\Lambda}} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{1}{nT} \sum_{t=1}^{T} \ddot{\boldsymbol{u}}' \boldsymbol{Z}_t' \boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}} \boldsymbol{Z}_t \ddot{\boldsymbol{u}} \right\} \\
&\geq a_{nT}^2 \mu_P \left(\frac{1}{nT} \boldsymbol{\mathcal{Z}}' (\boldsymbol{I}_T \otimes \boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}) \boldsymbol{\mathcal{Z}} \right) ||\ddot{\boldsymbol{u}}||_2^2 \ \geq a_{nT}^2 c ||\ddot{\boldsymbol{u}}||_2^2 > 0,
\end{aligned} \tag{D.12}
$$

where the last inequality holds as $n, T \to \infty$, because the matrix $\sum_{t=1}^{T} \boldsymbol{Z}_t' \boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}} \boldsymbol{Z}_t = \boldsymbol{\mathcal{Z}}'(\boldsymbol{I}_T \otimes \boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}) \boldsymbol{\mathcal{Z}}$ converges in probability to a positive definite matrix by Assumption 4.2. Thus, this matrix has real eigenvalues and is diagonalisable with orthogonal eigenvectors and, as such, by the same steps as in the proof of Lemma A.3(i) (see equation (E.5)), it is straightforward to show that this quadratic form is bounded from below by $\mu_P \left(\frac{1}{nT} \boldsymbol{\mathcal{Z}}'(\boldsymbol{I}_T \otimes \boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}) \boldsymbol{\mathcal{Z}} \right) ||\ddot{\boldsymbol{u}}||_2^2$, which in turn is bounded away from zero as $n, T \to \infty$. Next,

$$
\begin{aligned}
|k_2| &= \frac{2}{nT} a_{nT} \sum_{p=1}^{P} |\ddot{u}_p| |\text{tr}(\mathcal{Z}_p' \boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1} \boldsymbol{\varepsilon})| \leq \frac{2}{nT} \left(\sum_{p=1}^{P} |\ddot{u}_p|^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^{P} |\text{tr}(\mathcal{Z}_p' \boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1} \boldsymbol{\varepsilon})|^2 \right)^{\frac{1}{2}} \\
&= a_{nT} ||\ddot{\boldsymbol{u}}||_2 O_P \left(\sqrt{\frac{P}{nT}} \right),
\end{aligned} \tag{D.13}
$$

where the last line follows using Lemma A.2(v) and Markov's inequality. For term k_3,

$$
\begin{aligned}
|k_3| &= \left| \sup_{\dot{\boldsymbol{\Lambda}} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{2}{nT} a_{nT} \sum_{p=1}^{P} \ddot{u}_p \text{tr}(\mathcal{Z}_p' \boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}} \boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}) \right\} \right| \\
&\leq \frac{2(R+R^0)}{nT} a_{nT} \sum_{p=1}^{P} |\ddot{u}_p| ||\mathcal{Z}_p||_2 ||\boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1}||_2 ||\boldsymbol{\varepsilon}||_2 \\
&\leq \frac{2(R+R^0)}{nT} a_{nT} ||\ddot{\boldsymbol{u}}||_2 ||\boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1}||_2 ||\boldsymbol{\varepsilon}||_2 \left(\sum_{p=1}^{P} ||\mathcal{Z}_p||_2^2 \right)^{\frac{1}{2}} \\
&= a_{nT} ||\ddot{\boldsymbol{u}}||_2 O_P \left(\sqrt{\frac{P}{\min\{n, T\}}} \right),
\end{aligned} \tag{D.14}
$$

which follows because $||\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}}||_2 = 1$, since the maximum eigenvalue of any projection matrix is 1, $||\boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1}||_2 \leq \sqrt{||\boldsymbol{S}(\ddot{\boldsymbol{\rho}})||_1 ||\boldsymbol{S}(\ddot{\boldsymbol{\rho}})||_\infty} \sqrt{||\boldsymbol{S}^{-1}||_1 ||\boldsymbol{S}^{-1}||_\infty}$ and both $\boldsymbol{S}(\ddot{\boldsymbol{\rho}})$ and $\boldsymbol{S}$ are UB by Assumption 2.2, $||\boldsymbol{\varepsilon}||_2 = O_P \left(\frac{1}{\sqrt{\min\{n,T\}}} \right)$, and Lemma A.2(iv). Next,

$$
k_4 = \frac{\sigma_0^2}{n} \text{tr}((\boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1})' \boldsymbol{S}(\ddot{\boldsymbol{\rho}}) \boldsymbol{S}^{-1}) + O_P \left(\frac{1}{\sqrt{nT}} \right) \tag{D.15}
$$

by Lemma 9 in Yu et al. (2008). For the last term,

$$|k_5| = \sup_{\dot{\mathbf{\Lambda}} \in \mathbb{R}^{n\times(R+R^0)}} \left\{ \frac{1}{nT} \mathrm{tr}((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon})' \boldsymbol{P}_{\dot{\mathbf{\Lambda}}} \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}) \right\},$$

and thus

$$\begin{aligned} |k_5| &\leq \frac{(R+R^0)}{nT} ||\boldsymbol{P}_{\dot{\mathbf{\Lambda}}} \boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}||_2^2 \\ &\leq \frac{(R+R^0)}{nT} ||\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2^2 ||\boldsymbol{\varepsilon}||_2^2 = O_P\left(\frac{1}{\min\{n,T\}}\right) \end{aligned} \tag{D.16}$$

using the probability order of $||\boldsymbol{\varepsilon}||_2^2$, and the fact that the matrices $\boldsymbol{S}(\ddot{\boldsymbol{\rho}}), \boldsymbol{S}^{-1}$ are UB. Combining all the above gives

$$\begin{aligned} \mathcal{L}(\ddot{\boldsymbol{\theta}}) \leq \frac{1}{n}\log(\det(\boldsymbol{S}(\ddot{\boldsymbol{\rho}}))) - \frac{1}{2}\log\Bigg(& a_{nT}^2 c||\ddot{\boldsymbol{u}}||_2^2 + a_{nT}||\ddot{\boldsymbol{u}}||_2 O_P\left(\sqrt{\frac{P}{nT}}\right) + a_{nT}||\ddot{\boldsymbol{u}}||_2 O_P\left(\sqrt{\frac{P}{\min\{n,T\}}}\right) \\ & + \frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}) + O_P\left(\frac{1}{\sqrt{nT}}\right) + O_P\left(\frac{1}{\min\{n,T\}}\right)\Bigg) =: \bar{\mathcal{L}}(\ddot{\boldsymbol{\theta}}). \end{aligned} \tag{D.17}$$

Now, equation (D.2) is satisfied if $\mathcal{L}(\ddot{\boldsymbol{\theta}}) - \mathcal{L}(\boldsymbol{\theta}^0) < 0$ as $n, T \to \infty$. Since $\underline{\mathcal{L}}(\boldsymbol{\theta}^0) \leq \mathcal{L}(\boldsymbol{\theta}^0)$, and $\mathcal{L}(\ddot{\boldsymbol{\theta}}) \leq \bar{\mathcal{L}}(\ddot{\boldsymbol{\theta}})$, then equation (D.2) is equivalently satisfied if $\bar{\mathcal{L}}(\ddot{\boldsymbol{\theta}}) - \underline{\mathcal{L}}(\boldsymbol{\theta}^0) \leq 0$ as $n, T \to \infty$. Combining (D.10) and (D.17),

$$\begin{aligned} \bar{\mathcal{L}}(\ddot{\boldsymbol{\theta}}) - \underline{\mathcal{L}}(\boldsymbol{\theta}^0) \leq \frac{1}{2}\log\left(\sigma_0^2 \det((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}}\right) - \frac{1}{2}\log\Bigg(& a_{nT}^2 c||\ddot{\boldsymbol{u}}||_2^2 + a_{nT}||\ddot{\boldsymbol{u}}||_2 O_P\left(\sqrt{\frac{P}{nT}}\right) \\ & + a_{nT}||\ddot{\boldsymbol{u}}||_2 O_P\left(\sqrt{\frac{P}{\min\{n,T\}}}\right) + \frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'(\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1}) \\ & + O_P\left(\frac{1}{\sqrt{nT}}\right) + O_P\left(\frac{1}{\min\{n,T\}}\right)\Bigg), \end{aligned} \tag{D.18}$$

since

$$\begin{aligned} & \frac{1}{n}\log(\det(\boldsymbol{S}(\ddot{\boldsymbol{\rho}}))) - \frac{1}{n}\det(\boldsymbol{S}) + \frac{1}{2}\log\left(\sigma_0^2 + O_P\left(\frac{1}{\min\{n,T\}}\right)\right) \\ & = \frac{1}{2}\log\left(\sigma_0^2 \det((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}} + O_P\left(\frac{1}{\min\{n,T\}}\right)\right). \end{aligned} \tag{D.19}$$

Ignoring dominated terms, (D.18) becomes

$$\bar{\mathcal{L}}(\ddot{\boldsymbol{\theta}}) - \underline{\mathcal{L}}(\boldsymbol{\theta}^0) \leq \frac{1}{2}\log\left(\sigma_0^2 \det((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}} + O_P\left(\frac{1}{\min\{n,T\}}\right)\right)$$

$$-\frac{1}{2}\log\left(a_{nT}^2 c||\ddot{\boldsymbol{u}}||_2^2 + \mathrm{tr}((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})\right). \tag{D.20}$$

Recall that $c > 0$, and note that by Lemma A.1, $\sigma_0^2 \det((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}} \leq \frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\ddot{\boldsymbol{\rho}})\boldsymbol{S}^{-1})$. Then, by the monotonicity of the logarithm, as $n, T \to \infty$ and for sufficiently large d, the right-hand side of (D.20) is strictly negative. Therefore with probability approaching 1 there exists an a_{nT}-consistent local maximiser $\tilde{\boldsymbol{\theta}}_L$ of the unpenalised average likelihood function $\mathcal{L}(\boldsymbol{\theta})$. With the existence of a local maximiser established, consider next a global maximiser $\tilde{\boldsymbol{\theta}} := \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{L}(\boldsymbol{\theta})$. From (D.17), an upper bound for $\mathcal{L}(\tilde{\boldsymbol{\theta}})$ is given by

$$\begin{aligned}\mathcal{L}(\tilde{\boldsymbol{\theta}}) \leq \frac{1}{n}\log(\det(\boldsymbol{S}(\hat{\boldsymbol{\rho}}))) - \frac{1}{2}\log\Big(c||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2^2 + O_P(a_{nT})||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0|| + O_P(a_{nT}^2)\\ + \frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})\Big) =: \bar{\mathcal{L}}(\tilde{\boldsymbol{\theta}}).\end{aligned} \tag{D.21}$$

Since $\tilde{\boldsymbol{\theta}}$ is a global maximiser $\mathcal{L}(\boldsymbol{\theta}^0) \leq \mathcal{L}(\tilde{\boldsymbol{\theta}})$ and therefore $\underline{\mathcal{L}}(\boldsymbol{\theta}^0) \leq \bar{\mathcal{L}}(\tilde{\boldsymbol{\theta}})$. Combining this with (D.7) and (D.21) gives

$$\begin{aligned}&\frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\log\left(\sigma_0^2 + O_P\left(\frac{1}{\min\{n,T\}}\right)\right)\\ &\leq \frac{1}{n}\log(\det(\boldsymbol{S}(\hat{\boldsymbol{\rho}}))) - \frac{1}{2}\log\Big(c||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2^2 + O_P(a_{nT})||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0|| + O_P(a_{nT}^2)\\ &\quad + \frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})\Big).\end{aligned} \tag{D.22}$$

Multiplying both sides of (D.22) by -2, exponentiating, and then noticing that, by Lemma A.1, $\sigma_0^2\det((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}} \leq \frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})$, results in

$$0 \geq c||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2^2 + O_P(a_{nT})||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0|| + O_P(a_{nT}^2). \tag{D.23}$$

Completing the square, $0 \geq (\sqrt{c}||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2 + O_P(a_{nT}))^2 + O_P(a_{nT}^2)$, whereby it follows that $||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2 = O_P(a_{nT})$. Combined with the existence of a local maximiser, this result demonstrates the existence of an a_{nT}-consistent global maximiser of the unpenalised likelihood. Moving to the penalised average likelihood $\mathcal{Q}(\boldsymbol{\theta})$, and using the same notation $\ddot{\boldsymbol{u}}$ to denote $\ddot{\boldsymbol{u}} := \arg\sup_{||\boldsymbol{u}||_2=d}\{\mathcal{Q}(\boldsymbol{\theta}^0 + a_{nT}\boldsymbol{u})\}$,

$$\mathcal{Q}(\ddot{\boldsymbol{\theta}}) - \mathcal{Q}(\boldsymbol{\theta}^0) \leq \mathcal{L}(\ddot{\boldsymbol{\theta}}) - \mathcal{L}(\boldsymbol{\theta}^0) - \sum_{p=1}^{P_0}\varrho_p(\ddot{\theta}_p, \gamma_p, \zeta_p) + \sum_{p=1}^{P_0}\varrho_p(\theta_p^0, \gamma_p, \zeta_p). \tag{D.24}$$

For the penalty term,

$$\left|-\sum_{p=1}^{P_0}\varrho_p(\ddot{\theta}_p, \gamma_p, \zeta_p) + \sum_{p=1}^{P_0}\varrho_p(\theta_p^0, \gamma_p, \zeta_p)\right| = \left|\sum_{p=1}^{P_0}\gamma_p\frac{1}{|\theta_p^\dagger|^{\zeta_p}}|\theta_p^0 + a_{nT}\ddot{u}_p| - \sum_{p=1}^{P_0}\gamma_p\frac{1}{|\theta_p^\dagger|^{\zeta_p}}|\theta_p^0|\right|$$

$$\leq a_{nT} \sum_{p=1}^{P_0} \gamma_p \frac{1}{|\theta_p^\dagger|^{\zeta_p}} |\ddot{u}_p|,$$

where the last line follows from the triangle inequality. By the Cauchy-Schwarz inequality,

$$\begin{aligned} a_{nT} \sum_{p=1}^{P_0} \gamma_p \frac{1}{|\theta_p^\dagger|^{\zeta_p}} |\ddot{u}_p| &\leq a_{nT} \left(\sum_{p=1}^{P_0} \left(\frac{\gamma_p}{|\theta_p^\dagger|^{\zeta_p}} \right)^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^{P_0} |\ddot{u}_p|^2 \right)^{\frac{1}{2}} \\ &= a_{nT} \left(\sum_{p=1}^{P_0} \left(\frac{\gamma_p}{|\theta_p^\dagger|^{\zeta_p}} \right)^2 \right)^{\frac{1}{2}} ||\ddot{\boldsymbol{u}}||_2 \\ &\leq a_{nT} \left(P_0 \max_{1 \leq p \leq P^0} \left\{ \left(\frac{\gamma_p}{|\theta_p^\dagger|^{\zeta_p}} \right)^2 \right\} \right)^{\frac{1}{2}} ||\ddot{\boldsymbol{u}}||_2 = a_{nT} \sqrt{P_0} \frac{\gamma_{\bar{p}}}{|\theta_{\bar{p}}^\dagger|^{\zeta_{\bar{p}}}} ||\ddot{\boldsymbol{u}}||_2, \end{aligned} \tag{D.25}$$

with $\bar{p} := \arg\max_{1 \leq p \leq P^0} (\gamma_p |\theta_p^\dagger|^{-\zeta_p})^2$. Equation (D.25) can be rewritten as

$$a_{nT} \sqrt{\frac{P_0}{\min\{n,T\}}} \frac{\gamma_{\bar{p}} \sqrt{\min\{n,T\}}}{|\theta_{\bar{p}}^0|^{\zeta_{\bar{p}}}} \left(\frac{\theta_{\bar{p}}^\dagger}{\theta_{\bar{p}}^0} \right)^{-\zeta_{\bar{p}}} ||\ddot{\boldsymbol{u}}||_2. \tag{D.26}$$

Since the initial estimate $\boldsymbol{\theta}^\dagger$ satisfies $||\boldsymbol{\theta}^\dagger - \boldsymbol{\theta}^0||_2 = O_P(c_{nT}) = o_P(1)$, it follows that $|\theta_{\bar{p}}^\dagger / \theta_{\bar{p}}^0 - 1| \leq \frac{1}{|\theta_{\bar{p}}^0|} ||\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2 = o_P(1)$ which implies $\theta_{\bar{p}}^\dagger / \theta_{\bar{p}}^0 = O_P(1)$. Also, by Assumption 3.1 $\gamma_p \sqrt{\min\{n,T\}} |\theta_p^0|^{-\zeta_p} = O(1)$ for all $p = 1, ..., P^0$, and so

$$-\sum_{p=1}^{P_0} \varrho_p(\ddot{\theta}_p, \gamma_p, \zeta_p) + \sum_{p=1}^{P_0} \varrho_p(\theta_p^0, \gamma_p, \zeta_p) = O_P(a_{nT}^2) ||\ddot{\boldsymbol{u}}||_2, \tag{D.27}$$

whereby $\mathcal{Q}(\ddot{\boldsymbol{\theta}}) - \mathcal{Q}(\boldsymbol{\theta}^0) = \mathcal{L}(\ddot{\boldsymbol{\theta}}) - \mathcal{L}(\boldsymbol{\theta}^0) + O_P(a_{nT}^2) ||\ddot{\boldsymbol{u}}||_2$. It has already been established that, for large enough d, $\mathcal{L}(\ddot{\boldsymbol{\theta}}) - \mathcal{L}(\boldsymbol{\theta}^0)$ is strictly negative as $n, T \to \infty$, therefore it follows from equation (D.1) that there exists a local maximiser of the average penalised likelihood, $\hat{\boldsymbol{\theta}}_L$, in the interior of the ball $\{\boldsymbol{\theta}^0 + a_{nT} \ddot{\boldsymbol{u}} : ||\ddot{\boldsymbol{u}}||_2 \leq d\}$, such that $||\hat{\boldsymbol{\theta}}_L - \boldsymbol{\theta}^0||_2 = O_P(a_{nT})$. By the same steps used to derive (D.23), it can be shown that a global maximiser $\hat{\boldsymbol{\theta}}$ of the unpenalised likelihood must be a_{nT}-consistent, whereby both the existence and a_{nT}-consistency of the global maximum of the penalised likelihood is established. $\square$

Proof of Proposition 4. Let $\boldsymbol{\gamma}^0$ by some $\boldsymbol{\gamma}$ which satisfies Assumptions 3.1 and 5. From Propositions 1 and 2, with probability approaching 1, the true model is selected, in which case $\boldsymbol{\gamma}^0 \in \Gamma^0$. Moreover, since under $\boldsymbol{\gamma}^0$, $||\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2 = O_P(a_{nT})$, it follows that $\hat{\sigma}^2(\boldsymbol{\gamma}^0) = \sigma_0^2 + o_P(1)$ using Lemmas F.1(x) and F.3(v). Hence,

$$\begin{aligned}\mathrm{IC}^*(\boldsymbol{\gamma}^0) &= \hat{\sigma}^2(\boldsymbol{\gamma}^0) + \varrho_\rho Q^0 + \varrho_\beta K^0 \\ &= \sigma_0^2 + o_P(1), \end{aligned} \tag{D.28}$$

as $Q^0\varrho_\rho, K^0\varrho_\beta \to 0$ by Assumption 8.1. Now, consider some $\boldsymbol{\gamma} \in \Gamma^+$ that produces an overfitted model. It is shown that, as $n, T \to \infty$,

$$\Pr\left(\inf_{\boldsymbol{\gamma}\in\Gamma^+} \mathrm{IC}^*(\boldsymbol{\gamma}) > \mathrm{IC}^*(\boldsymbol{\gamma}^0)\right) \to 1. \tag{D.29}$$

Recalling that $\mathrm{IC}^*(\boldsymbol{\gamma}) \coloneqq \hat{\sigma}^2(\boldsymbol{\gamma}) + \varrho_\rho|\mathcal{S}_\rho(\boldsymbol{\gamma})| + \varrho_\beta|\mathcal{S}_\beta(\boldsymbol{\gamma})|$, (D.29) is equivalent to

$$\Pr\Big(\inf_{\boldsymbol{\gamma}\in\Gamma^+}\{\hat{\sigma}^2(\boldsymbol{\gamma}) + \varrho_\rho|\mathcal{S}_\rho(\boldsymbol{\gamma})| + \varrho_\beta|\mathcal{S}_\beta(\boldsymbol{\gamma})|\} > \hat{\sigma}^2(\boldsymbol{\gamma}^0) + \varrho_\rho Q^0 + \varrho_\beta K^0\Big) \to 1. \tag{D.30}$$

Let $\boldsymbol{\gamma}^+ \coloneqq \arg\inf_{\boldsymbol{\gamma}\in\Gamma^+} \mathrm{IC}^*(\boldsymbol{\gamma})$. Then (D.30) gives

$$\begin{aligned}\Pr\Big(&(\sqrt{Q}a_{nT})^{-1}(\hat{\sigma}^2(\boldsymbol{\gamma}^+) - \hat{\sigma}^2(\boldsymbol{\gamma}^0)) + (\sqrt{Q}a_{nT})^{-1}\varrho_\rho(|\mathcal{S}_\rho(\boldsymbol{\gamma}^+)| - Q^0) \\ &+(\sqrt{Q}a_{nT})^{-1}\varrho_\beta(|\mathcal{S}_\beta(\boldsymbol{\gamma}^+)| - K^0) > 0\Big) \to 1. \end{aligned} \tag{D.31}$$

An overfitted model does not exclude any relevant variables and therefore it is straightforward to show, using the same steps as in the proof of Proposition 1, that the estimator under $\boldsymbol{\gamma}^+$ - call this $\hat{\boldsymbol{\theta}}_+$ - satisfies $||\hat{\boldsymbol{\theta}}_+ - \boldsymbol{\theta}^0||_2 = O_P(a_{nT})$. Moreover, it is also possible to derive an expression for $\hat{\sigma}^2(\boldsymbol{\gamma}^+)$ analogous to that derived in Lemma F.3(v). Given this, it can be seen that $\hat{\sigma}^2(\boldsymbol{\gamma}^+) - \hat{\sigma}^2(\boldsymbol{\gamma}^0) = O_P(\sqrt{Q}a_{nT})$, and so $(\sqrt{Q}a_{nT})^{-1}(\hat{\sigma}^2(\boldsymbol{\gamma}^+) - \hat{\sigma}^2(\boldsymbol{\gamma}^0)) = O_P(1)$. By Assumption 8.1, $(\sqrt{Q}a_{nT})^{-1}\varrho_\rho, (\sqrt{Q}a_{nT})^{-1}\varrho_\beta \to \infty$, and since either $|\mathcal{S}_\rho(\boldsymbol{\gamma}^+)| - K^0 > 0$ or $|\mathcal{S}_\beta(\boldsymbol{\gamma}^+)| - Q^0 > 0$, or both, then (D.31) holds as $n, T \to \infty$. Finally, in the case of an underfitted model, one of either $\mathcal{S}_\rho(\boldsymbol{\gamma}) \not\supset \mathcal{S}_{T,\rho}$ or $\mathcal{S}_\beta(\boldsymbol{\gamma}) \not\supset \mathcal{S}_{T,\beta}$ must be true. Then,

$$\inf_{\boldsymbol{\gamma}\in\Gamma^-} \mathrm{IC}^*(\boldsymbol{\gamma}) > \inf_{\boldsymbol{\gamma}\in\Gamma^-} \hat{\sigma}^2(\boldsymbol{\gamma}) \xrightarrow{p} \sigma_-^2 > \sigma_0^2, \tag{D.32}$$

using Assumption 8.2. Hence $\Pr\left(\inf_{\boldsymbol{\gamma}\in\Gamma^-} \mathrm{IC}^*(\boldsymbol{\gamma}) > \mathrm{IC}^*(\boldsymbol{\gamma}^0)\right) \to 1$. Combined, (D.29) and (D.32) establish the result.

□

E Proofs of Lemmas A.1–A.3

This Appendix provides proofs of Lemmas A.1–A.3.

Proof of Lemma A.1. Since the trace of a matrix is the sum of its eigenvalues, and the determinant of a matrix is the product of its eigenvalues, it follows that for any positive definite matrix $\boldsymbol{B}$, $\det(\boldsymbol{B})^{\frac{1}{n}} \leq \frac{1}{n}\mathrm{tr}(\boldsymbol{B})$ by the inequality of arithmetic-geometric means. This inequality is satisfied with equality if and only if of all of the eigenvalues of $\boldsymbol{B}$ are the same, in which case $\boldsymbol{B} = c\boldsymbol{I}_n$, for some constant c. Since $\boldsymbol{B}$ is positive definite, all of its eigenvalues are positive, whereby c must be strictly positive. □

Proof of Lemma A.2(i). Recall $\boldsymbol{S}(\boldsymbol{\rho}) \coloneqq \boldsymbol{I}_n - \sum_{q=1}^{Q} \rho_q \boldsymbol{W}_q$. Then $\boldsymbol{S}(\boldsymbol{\rho}) + \sum_{q=1}^{Q} \rho_q \boldsymbol{W}_q = \boldsymbol{I}_n$, and therefore $\boldsymbol{I}_n + \sum_{q=1}^{Q} \rho_q \boldsymbol{G}_q = \boldsymbol{S}^{-1}(\boldsymbol{\rho})$. Now,

$$
\begin{aligned}
\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1} &= \left(\boldsymbol{I}_n - \sum_{q=1}^{Q} \rho_q \boldsymbol{W}_q\right)\left(\boldsymbol{I}_n + \sum_{q=1}^{Q} \rho_q^0 \boldsymbol{G}_q\right) \\
&= \boldsymbol{I}_n + \sum_{q=1}^{Q} \rho_q^0 \boldsymbol{G}_q \boldsymbol{S}^{-1} - \sum_{q=1}^{Q} \rho_q \boldsymbol{W}_q \left(\boldsymbol{I}_n + \sum_{q=1}^{Q} \rho_q^0 \boldsymbol{G}_q\right) \\
&= \boldsymbol{I}_n + \sum_{q=1}^{Q} (\rho_q^0 - \rho_q) \boldsymbol{G}_q.
\end{aligned}
$$

□

Proof of Lemma A.2(ii). There are four types of covariate to consider: $\boldsymbol{\mathcal{X}}_\kappa^*$, $\boldsymbol{\mathcal{Y}}_{-1}$, $\boldsymbol{W}_q \boldsymbol{\mathcal{Y}}_{-1}$ and $\sum_{k=1}^{K^0} \beta^0 \boldsymbol{G}_q \boldsymbol{\mathcal{X}}_k$, for some κ and q. First for the κ-th exogenous covariate,

$$
\mathbb{E}\left[||\boldsymbol{\mathcal{X}}_\kappa^*||_2^2\right] \leq \mathbb{E}\left[||\boldsymbol{\mathcal{X}}_\kappa^*||_F^2\right] = \sum_{i=1}^{n}\sum_{t=1}^{T} \mathbb{E}\left[(x_{\kappa it}^*)^2\right].
$$

By Assumption 2.4 the fourth moment of $x_{\kappa it}^*$ is uniformly bounded, and therefore $\mathbb{E}[||\boldsymbol{\mathcal{X}}_\kappa^*||_2^2] = O(nT)$. By Markov's inequality,

$$
\Pr\left(||\boldsymbol{\mathcal{X}}_\kappa^*||_F^2 > \frac{cnT}{\epsilon}\right) \leq \frac{\epsilon}{cnT}\mathbb{E}\left[||\boldsymbol{\mathcal{X}}_\kappa^*||_F^2\right] < \epsilon,
$$

for all $\varepsilon > 0$ and so $||\boldsymbol{\mathcal{X}}_\kappa^*||_F = O_P(\sqrt{nT})$. Note also that for any variables generated as $\boldsymbol{W}_q \boldsymbol{\mathcal{X}}_\kappa^*$,

$$
\begin{aligned}
\mathbb{E}\left[((\boldsymbol{W}_q \boldsymbol{\mathcal{X}}_\kappa^*)_{it})^2\right] &= \mathbb{E}\left[\left(\sum_{j=1}^{n} (\boldsymbol{W}_q)_{ij} x_{\kappa jt}^*\right)^2\right] = \sum_{j=1}^{n}\sum_{j'=1}^{n} (\boldsymbol{W}_q)_{ij} (\boldsymbol{W}_q)_{ij'} \mathbb{E}\left[x_{\kappa jt}^* x_{\kappa j't}^*\right] \\
&\leq \left(\sum_{j=1}^{n} |(\boldsymbol{W}_q)_{ij}|\right)\left(\sum_{j=1}^{n} |(\boldsymbol{W}_q)_{ij'}|\right)\left(\mathbb{E}\left[(x_{\kappa jt}^*)^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[(x_{\kappa j't}^*)^2\right]\right)^{\frac{1}{2}}
\end{aligned}
$$

$$\leq c_1,$$

because the fourth moment of $x^*_{\kappa it}$ is uniformly bounded, and the weights matrix $\boldsymbol{W}_q$ is UB uniformly over q. It then follows from Markov's inequality that $||\boldsymbol{W}_q \boldsymbol{\mathcal{X}}^*_\kappa||_F = O_P(\sqrt{nT})$. Next, consider the elements of $\boldsymbol{W}_q \boldsymbol{\mathcal{Y}}_{-1}$. Recursive substitution gives

$$\boldsymbol{W}_q \boldsymbol{y}_t = \sum_{h=0}^{\infty} \boldsymbol{W}_q \boldsymbol{A}^h \boldsymbol{S}^{-1} \left(\sum_{\kappa=1}^{K^*} \delta^0_\kappa \boldsymbol{x}^*_{\kappa t-h} + \boldsymbol{\Lambda}^0 \boldsymbol{f}^0_{t-h} + \boldsymbol{\varepsilon}_{t-h} \right),$$

so that, for the (i,t)-th element,

$$\begin{aligned}
(\boldsymbol{W}_q \boldsymbol{\mathcal{Y}}_{-1})_{it} &= \sum_{h=0}^{\infty} \sum_{j=1}^{n} \sum_{\kappa=1}^{K^*} (\boldsymbol{\mathcal{A}}(h))_{ij} \delta^0_\kappa x^*_{\kappa jt-h} + \sum_{h=0}^{\infty} \sum_{j=1}^{n} \sum_{r=1}^{R^0} (\boldsymbol{\mathcal{A}}(h))_{ij} \lambda^0_{rj} f^0_{rt-h} + \sum_{h=0}^{\infty} \sum_{j=1}^{n} (\boldsymbol{\mathcal{A}}(h))_{ij} \varepsilon_{jt-h} \\
&=: l_1 + l_2 + l_3,
\end{aligned}$$

where $\boldsymbol{\mathcal{A}}(h) := \boldsymbol{W}_q \boldsymbol{A}^h \boldsymbol{S}^{-1}$. First,

$$\begin{aligned}
\mathbb{E}[l_1^2] &= \sum_{h=0}^{\infty} \sum_{j=1}^{n} \sum_{\kappa=1}^{K^*} \sum_{h'=0}^{\infty} \sum_{j'=1}^{n} \sum_{\kappa'=1}^{K^*} (\boldsymbol{\mathcal{A}}(h))_{ij} (\boldsymbol{\mathcal{A}}(h'))_{ij'} \delta^0_\kappa \delta^0_{\kappa'} \mathbb{E}\left[x^*_{\kappa jt-h} x^*_{\kappa' j' t-h'}\right] \\
&\leq \sum_{h=0}^{\infty} \sum_{j=1}^{n} \sum_{\kappa=1}^{K^*} \sum_{h'=0}^{\infty} \sum_{j'=1}^{n} \sum_{\kappa'=1}^{K^*} |(\boldsymbol{\mathcal{A}}(h))_{ij}| |(\boldsymbol{\mathcal{A}}(h'))_{ij'}| |\delta^0_\kappa| |\delta^0_{\kappa'}| \left(\mathbb{E}\left[(x^*_{\kappa jt-h})^2\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[(x^*_{\kappa j' t-h'})^2\right]\right)^{\frac{1}{2}} \\
&\leq c_1 \left(\sum_{h=0}^{\infty} \sum_{j=1}^{n} |(\boldsymbol{\mathcal{A}}(h))_{ij}| \right)^2 \left(\sum_{\kappa=1}^{K^*} |\delta^0_\kappa| \right)^2 \leq c_2,
\end{aligned}$$

using Assumption 2.5, the assumption that has a $x^*_{\kappa it}$ uniformly bounded fourth moment, and also that because

$$\left|\left| \sum_{h=0}^{\infty} \boldsymbol{\mathcal{A}}(h) \right|\right|_\infty \leq ||\boldsymbol{W}_q||_\infty ||\boldsymbol{S}^{-1}||_\infty \left|\left| \sum_{h=0}^{\infty} \boldsymbol{A}^h \right|\right|_\infty, \tag{E.1}$$

then $\sum_{j=1}^n | (\sum_{h=0}^\infty (\boldsymbol{\mathcal{A}}(h))_{ij}) | \leq || \sum_{h=0}^\infty (\boldsymbol{\mathcal{A}}(h))_{ij} ||_\infty < c_3$ by Assumption 2.3. For l_2,

$$\begin{aligned}
\mathbb{E}[l_2^2] &= \sum_{h=0}^{\infty} \sum_{j=1}^{n} \sum_{r=1}^{R^0} \sum_{h'=0}^{\infty} \sum_{j'=1}^{n} \sum_{r'=1}^{R^0} (\mathcal{A}(h))_{ij} (\mathcal{A}(h'))_{ij'} \mathbb{E}\left[\lambda^0_{rj} \lambda^0_{r'j'} f^0_{rt-h} f^0_{r't-h'}\right] \\
&\leq \sum_{h=0}^{\infty} \sum_{j=1}^{n} \sum_{r=1}^{R^0} \sum_{h'=0}^{\infty} \sum_{j'=1}^{n} \sum_{r'=1}^{R^0} |(\mathcal{A}(h))_{ij}| |(\mathcal{A}(h'))_{ij'}| \left(\mathbb{E}\left[(\lambda^0_{rj})^4\right]\right)^{\frac{1}{4}} \left(\mathbb{E}\left[(\lambda^0_{r'j'})^4\right]\right)^{\frac{1}{4}} \\
&\quad \times \left(\mathbb{E}\left[(f^0_{rt-h})^4\right]\right)^{\frac{1}{4}} \left(\mathbb{E}\left[(f^0_{r't-h'})^4\right]\right)^{\frac{1}{4}}
\end{aligned}$$

$$\leq R^{0^2} \left(\sum_{h=0}^{\infty} \sum_{j=1}^{n} |(\mathcal{A}(h))_{ij}| \right)^2 c_4 \leq c_5,$$

using (E.1). Similarly, $\mathbb{E}[l_3^2] \leq c_6$ using Assumption 1.1. As such it is straightforward to show that $\mathbb{E}\left[(\boldsymbol{W}_q \boldsymbol{\mathcal{Y}}_{-1})_{it}^2\right]$ is uniformly bounded across i and t, from whence it follows that $||\boldsymbol{W}_q \boldsymbol{\mathcal{Y}}_{-1}||_F = O_P(\sqrt{nT})$. The same steps also establish $||\boldsymbol{\mathcal{Y}}_{-1}||_F = O_P(\sqrt{nT})$ by replacing $\boldsymbol{W}_q$ with an identity matrix, which is trivially UB. Finally consider covariates of the form $\sum_{k=1}^{K^0} \beta_k^0 \boldsymbol{G}_q \boldsymbol{\mathcal{X}}_k$. It has been demonstrated above that the elements of and $\boldsymbol{W}_q \boldsymbol{\mathcal{Y}}_{-1}$ have uniformly bounded second moments. As such,

$$\mathbb{E}\left[\left(\sum_{k=1}^{K^0} \beta_k^0 (\boldsymbol{G}_q \boldsymbol{\mathcal{X}}_k)_{it} \right)^2 \right] = \sum_{k=1}^{K^0} \sum_{k'=1}^{K^0} \sum_{j=1}^{n} \sum_{j'=1}^{n} \beta_k^0 \beta_{k'}^0 (\boldsymbol{G}_q)_{ij} (\boldsymbol{G}_q)_{ij'} \mathbb{E}\left[x_{kit} x_{k'it}\right] \leq c_7, \quad \text{(E.2)}$$

using Assumptions 2.2–2.5 which gives the result $|| \sum_{k=1}^{K^0} \beta_k^0 \boldsymbol{G}_q \boldsymbol{\mathcal{X}}_k ||_F = O_P(\sqrt{nT})$. □

Proof of Lemma A.2(iii). Follows similar steps to the first part of the proof of Lemma A.2(ii) using Assumption 2.7. □

Proof of Lemma A.2(iv). Follows from Lemma A.2(ii). □

Proof of Lemma A.2(v). In the proof of Lemma 3 in Shi and Lee (2017) it is established that $\mathbb{E}\left[(\mathrm{tr}(\boldsymbol{\mathcal{Z}}_p' \boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}))^2\right] = O(nT)$ for each p. Minor modification to that lemma yields the result. □

Proof of Lemma A.2(vi). Using Assumption 1.1, $\mathbb{E}\left[||\boldsymbol{\varepsilon}||_F^2\right] = \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i=1}^{n} \varepsilon_{it}^2\right] = nT\sigma_0^2 = O(nT)$. The proof is completed using Markov's inequality. □

Proof of Lemma A.2(vii).

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} ||\boldsymbol{X}_t \boldsymbol{\beta}^0||_2^2 \right] &= \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{k=1}^{K^0} \sum_{k'=1}^{K^0} \beta_k^0 \beta_{k'}^0 \mathbb{E}\left[x_{kit} x_{k'it}\right] \\
&\leq \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{k=1}^{K^0} \sum_{k'=1}^{K^0} |\beta_k^0||\beta_{k'}^0| \left(\mathbb{E}\left[x_{kit}^2\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[x_{k'it}^2\right]\right)^{\frac{1}{2}} \leq cnT,
\end{aligned}$$

using Assumption 2.5, and, as is shown in the proof of Lemma A.2(ii), the second moments of x_{kit} are uniformly bounded. □

Proof of Lemma A.2(viii). Using Lemma A.2(i), $||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n||_2 \leq \sum_{q=1}^{Q} |\rho_q^0 - \hat{\rho}_q|||\boldsymbol{W}_q\boldsymbol{S}^{-1}||_2$. Now,

$$
\begin{aligned}
\sum_{q=1}^{Q} |\rho_q^0 - \hat{\rho}_q|||\boldsymbol{W}_q\boldsymbol{S}^{-1}||_2 &\leq ||\boldsymbol{S}^{-1}||_2 \sum_{q=1}^{Q} |\rho_q^0 - \hat{\rho}_q|||\boldsymbol{W}_q||_2 \qquad \text{(E.3)}\\
&\leq ||\boldsymbol{S}^{-1}||_2 \left(\sum_{q=1}^{Q} |\rho_q^0 - \hat{\rho}_q|^2\right)^{\frac{1}{2}} \left(\sum_{q=1}^{Q} ||\boldsymbol{W}_q||_2^2\right)^{\frac{1}{2}}\\
&= ||\boldsymbol{S}^{-1}||_2 ||\boldsymbol{\rho}^0 - \hat{\boldsymbol{\rho}}||_2 \sqrt{Q} \sqrt{\left(\max_{1\leq q\leq Q} ||\boldsymbol{W}_q||_2^2\right)} = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)
\end{aligned}
$$

using Assumption 2.2 and $||\boldsymbol{\rho}^0 - \hat{\boldsymbol{\rho}}||_2 \leq ||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2$. □

Proof of Lemma A.3(i).

$$
\begin{aligned}
\left(\frac{1}{nT}\sum_{t=1}^{T} ||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}} &= \left(\frac{1}{nT}\sum_{t=1}^{T} \operatorname{tr}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\boldsymbol{Z}_t')\right)^{\frac{1}{2}}\\
&= \left((\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{Z}_t'\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\right)^{\frac{1}{2}}\\
&\leq \left(||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2 \mu_1\left(\frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{Z}_t'\boldsymbol{Z}_t\right)\right)^{\frac{1}{2}}\\
&= \left(||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2 \mu_1\left(\boldsymbol{\mathcal{H}}_2\right)\right)^{\frac{1}{2}} \leq c||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2, \qquad \text{(E.4)}
\end{aligned}
$$

where the last step uses Assumption 4.2 and the second to last comes from the following: since the matrix $\boldsymbol{\mathcal{H}}_2$ is symmetric and positive definite, it possess positive eigenvalues and orthogonal eigenvectors. Consider the eigendecomposition $\boldsymbol{\mathcal{H}}_2 = \boldsymbol{\mathcal{U}}'\boldsymbol{\Pi}\boldsymbol{\mathcal{U}}$,

$$
\begin{aligned}
(\boldsymbol{\theta}^0 - \boldsymbol{\theta})'\boldsymbol{\mathcal{H}}_2(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) &= (\boldsymbol{\theta}^0 - \boldsymbol{\theta})'\boldsymbol{\mathcal{U}}'\boldsymbol{\Pi}\boldsymbol{\mathcal{U}}(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) = ||\boldsymbol{\Pi}^{\frac{1}{2}}\boldsymbol{\mathcal{U}}(\boldsymbol{\theta}^0 - \boldsymbol{\theta})||_2^2\\
&\leq ||\boldsymbol{\Pi}^{\frac{1}{2}}||_2^2||\boldsymbol{\mathcal{U}}||_2^2||\boldsymbol{\theta}^0 - \boldsymbol{\theta}||_2^2 = \mu_1(\boldsymbol{\mathcal{H}}_2)||\boldsymbol{\theta}^0 - \boldsymbol{\theta}||_2^2. \qquad \text{(E.5)}
\end{aligned}
$$

□

Proof of Lemma A.3(ii). Recall that $\hat{\sigma}^2(\boldsymbol{\theta}, \boldsymbol{\Lambda}) := \frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{e}_t'\boldsymbol{M}_{\boldsymbol{\Lambda}}\boldsymbol{e}_t$, where $\boldsymbol{e}_t := \boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{y}_t - \boldsymbol{X}_t\boldsymbol{\beta}$. Evaluating at $\hat{\boldsymbol{\theta}}$, and substituting in the true DGP yields

$$
\hat{\sigma}^2(\hat{\boldsymbol{\theta}}, \boldsymbol{\Lambda}) = \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{X}_t\boldsymbol{\beta}^0 + \boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t - \boldsymbol{X}_t\hat{\boldsymbol{\beta}})'\boldsymbol{M}_{\boldsymbol{\Lambda}}
$$

$$(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{X}_t\boldsymbol{\beta}^0+\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0+\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t-\boldsymbol{X}_t\hat{\boldsymbol{\beta}}). \tag{E.6}$$

Using Lemma A.2(i), $\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{X}_t\boldsymbol{\beta}^0-\boldsymbol{X}_t\hat{\boldsymbol{\beta}}=\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})$. Applying this, and expanding (E.6) gives

$$\begin{aligned}
\hat{\sigma}^2(\hat{\boldsymbol{\theta}},\boldsymbol{\Lambda})=&\frac{1}{nT}\sum_{t=1}^T(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\boldsymbol{M}_{\boldsymbol{\Lambda}}\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})+\frac{2}{nT}\sum_{t=1}^T(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\boldsymbol{M}_{\boldsymbol{\Lambda}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\\
&+\frac{2}{nT}\sum_{t=1}^T(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\boldsymbol{M}_{\boldsymbol{\Lambda}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t+\frac{2}{nT}\sum_{t=1}^T(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\boldsymbol{\Lambda}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t\\
&+\frac{1}{nT}\sum_{t=1}^T(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\boldsymbol{\Lambda}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t+\frac{1}{nT}\sum_{t=1}^T(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\boldsymbol{\Lambda}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\\
=:&\ l_1,...,l_6. \end{aligned}\tag{E.7}$$

The probability order of terms l_1,l_3,l_5 is established following the same steps as those used for terms $k_1,...,k_5$ in the proof of Proposition 1. For the remaining terms,

$$\begin{aligned}
l_2&\leq\frac{2}{nT}\sum_{t=1}^T||\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})||_2||\boldsymbol{M}_{\boldsymbol{\Lambda}}||_2||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2||\boldsymbol{\Lambda}^0||_2||\boldsymbol{f}_t^0||_2\\
&\leq\frac{2}{\sqrt{nT}}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2\left(\frac{1}{nT}\sum_{t=1}^T||\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^T||\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}}\\
&=\frac{2}{\sqrt{nT}}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{F}^0||_F||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2\left(\frac{1}{nT}\sum_{t=1}^T||\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\\
&=\frac{1}{\sqrt{nT}}O_P(\sqrt{n})O_P(\sqrt{T})O_P(a_{nT})=O_P(a_{nT})
\end{aligned}\tag{E.8}$$

using Lemmas A.2(iii), A.3(i), Proposition 1 and noting that $||\boldsymbol{M}_{\boldsymbol{\Lambda}}||_2=1$. Next

$$\begin{aligned}
l_4&\leq\frac{2}{nT}R^0||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{M}_{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Lambda}^0||_2||\boldsymbol{F}^0||_2||\boldsymbol{\varepsilon}||_2\\
&=\frac{1}{nT}O_P(\sqrt{n})O_P(\sqrt{T})O_P(\sqrt{\max\{n,T\}})=O_P\left(\sqrt{\frac{1}{\min\{n,T\}}}\right)
\end{aligned}\tag{E.9}$$

using similar steps to those for l_2. For l_6,

$$\begin{aligned}
l_6&\leq\frac{1}{nT}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2^2||\boldsymbol{M}_{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Lambda}^0||_2^2\sum_{t=1}^T||\boldsymbol{f}_t^0||_2||\boldsymbol{f}_t^0||_2\\
&\leq\frac{1}{nT}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2^2||\boldsymbol{M}_{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Lambda}^0||_2^2\left(\sum_{t=1}^T||\boldsymbol{f}_t^0||_2\right)
\end{aligned}$$

$$= \frac{1}{nT}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2^2||\boldsymbol{M}_{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{F}^0||_F^2$$
$$= \frac{1}{nT}O_P(n)O_P(T) = O_P(1), \tag{E.10}$$

using Lemma A.2(iii) and because $||\boldsymbol{M}_{\boldsymbol{\Lambda}}||_2 = 1$. Note also that since the projection matrix $\boldsymbol{M}_{\boldsymbol{\Lambda}}$ is positive semidefinite, the quadratic form $l_6 \geq 0$. Combining all the above results, and ignoring dominated terms,

$$\hat{\sigma}^2(\hat{\boldsymbol{\theta}}, \boldsymbol{\Lambda}) = \frac{\sigma_0^2}{n}\text{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}) + O_P(1). \tag{E.11}$$

Now, by Lemma A.1 $\sigma_0^2\text{det}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}} \leq \frac{\sigma_0^2}{n}\text{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})$ and therefore a lower bound on (E.11) can be found where the above inequality is satisfied with equality. This occurs when $(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} = c_1\boldsymbol{I}_n$, with $c_1 > 0$. Therefore

$$0 < \frac{\sigma_0^2}{n}\text{tr}(c_1\boldsymbol{I}_n) = c_1\sigma_0^2 \leq \frac{\sigma_0^2}{n}\text{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}).$$

Next, recall that $|\text{tr}(\boldsymbol{B})| \leq \text{rank}(\boldsymbol{B})||\boldsymbol{B}||_2$. The $n \times n$ matrix $(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}$ can have rank no larger than n and so,

$$\frac{\sigma_0^2}{n}\text{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}) \leq \sigma_0^2||(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2 \leq \sigma_0^2||\boldsymbol{S}^{-1}||_2^2||\boldsymbol{S}(\hat{\boldsymbol{\rho}})||_2^2 = O(1),$$

since the matrices $\boldsymbol{S}^{-1}$ and $\boldsymbol{S}(\hat{\boldsymbol{\rho}})$ are UB. As a result $0 < \frac{\sigma_0^2}{n}\text{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}) = O(1)$. It then follows that $\hat{\sigma}^{-2}(\hat{\boldsymbol{\theta}}, \boldsymbol{\Lambda}) = O_P(1)$ and is strictly positive for large enough n, T. □

F Proofs of Lemmas A.4–A.6

This Appendix provides proofs of Lemmas A.4–A.6 for which some intermediary results are required in the form of Lemmas F.1–F.4. The proofs of these intermediary results can be found in Section G. Since Lemmas A.4–A.6 are used to prove Theorem 1 and Proposition 3, both of which only concern the correct model, for notational convenience it is assumed in this section that all of the regressors are relevant, i.e. $K = K^0, Q = Q^0$, $P = P^0$, $\boldsymbol{\mathcal{Z}}_{(1)} = \boldsymbol{\mathcal{Z}}$.

Lemma F.1 *Under Assumptions* 1–2,

(i) $||\hat{\boldsymbol{\Lambda}}||_2 = \sqrt{n}$, $||\hat{\boldsymbol{\Lambda}}||_F = \sqrt{Rn}$;

(ii) $||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'||_2 = O_P(\sqrt{nT})$, $||\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}||_F = O_P(\sqrt{nT})$, $||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}\boldsymbol{\Lambda}^0||_2 = O_P(\sqrt{nT})$;

(iii) $||(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}), ..., \text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}}))||_2 = O(\sqrt{Qn})$;

(iv) $||(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\varepsilon}),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\varepsilon}))||_2 = O_P(\sqrt{QnT})$;

(v) $\mathbb{E}[\sum_{p=1}^P ||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2] = O(P\max\{n,T\})$, *where* $\bar{\mathfrak{Z}}_p := \mathbb{E}_{\mathcal{D}}[\mathfrak{Z}_p]$;

(vi) $\mathbb{E}[\sum_{p=1}^P ||\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_F^2] = O(Pn^2T)$;

(vii) $\mathbb{E}[\sum_{p=1}^P ||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\bar{\mathfrak{Z}}_p||_F^2] = O(PnT^2)$, $\mathbb{E}[\sum_{p=1}^P ||\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_F^2] = O(Pn^2T)$, $\mathbb{E}[\sum_{p=1}^P\sum_{p'=1}^P ||\bar{\mathfrak{Z}}_p'\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_{p'}'||_F^2] = O(P^2n^2T^2)$;

(viii) $||\frac{1}{T}\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\boldsymbol{\Lambda}^{0'}\boldsymbol{\Lambda}^0||_F = O_P(\frac{n}{\sqrt{T}})$, $||\frac{1}{n}\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0 - \sigma_0^2\boldsymbol{F}^{0'}\boldsymbol{F}^0||_F = O_P(\frac{T}{\sqrt{n}})$;

(ix) $\mathbb{E}[\sum_{p=1}^P ||\frac{1}{T}\bar{\mathfrak{Z}}_p'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\bar{\mathfrak{Z}}_p'\boldsymbol{\Lambda}^0||_2^2] = O(Pn^2)$, $\mathbb{E}[\sum_{p=1}^P ||\frac{1}{n}\bar{\mathfrak{Z}}_p\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0 - \sigma_0^2\bar{\mathfrak{Z}}_p\boldsymbol{F}^0||_2^2] = O(PT^2)$;

(x) $\frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) = \sigma_0^2 + O_P\left(\frac{1}{\min\{n,T\}}\right)$.

Lemma F.2 *Under Assumptions* 1–6,

(i) $\frac{1}{\sqrt{n}}||\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2$ *and* $\frac{1}{\sqrt{T}}||\hat{\boldsymbol{F}}' - \boldsymbol{H}^{*-1}\boldsymbol{F}^{0'}||_2$ *are* $O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right)$, *where* $\boldsymbol{H}^* := \frac{1}{nT}\boldsymbol{F}^{0'}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Pi}^{-1}$ *and* $\hat{\boldsymbol{F}} := \frac{1}{n}\sum_{t=1}^T \hat{\boldsymbol{e}}_t'\hat{\boldsymbol{\Lambda}}$, *with* $\boldsymbol{\Pi}$ *being a diagonal* $R\times R$ *matrix containing the largest* R *eigenvalues of* $\frac{1}{nT}\sum_{t=1}^T \hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t'$ *along its diagonal and* $\hat{\boldsymbol{e}}_t := \boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{y}_t - \boldsymbol{X}_t\hat{\boldsymbol{\beta}}$;

(ii) *The matrix* $\frac{1}{n}\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}$ *converges in probability to an invertible matrix*;

(iii) $\frac{1}{nT}||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}||_2 = O_P\left(\frac{1}{\sqrt{nT}}\right) + O_P\left(\frac{1}{T}\right) + O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2}{\sqrt{T}}\right) + O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{\sqrt{nT},T\}}\right)$;

(iv) $-\frac{1}{n^2T^2}\sum_{t=1}^T\sum_{\tau=1}^T \boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 = \boldsymbol{O}_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right)$;

(v) $||\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2$ *and* $||\boldsymbol{P}_{\boldsymbol{F}^0} - \boldsymbol{P}_{\hat{\boldsymbol{F}}}||_2$ *are* $O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right)$;

(vi) $||\hat{\boldsymbol{F}}||_2 = O_P(\sqrt{T})$;

(vii) $\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon}) = \frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) + \boldsymbol{o}_P(1)$;

(viii) $\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) = \boldsymbol{O}_P(\sqrt{P})$;

(ix) $\frac{1}{nT}\mathbb{E}\left[||\mathbb{Z}'\mathbb{Z} - \boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}||_2\right] = o(1)$.

Lemma F.3 *Under Assumptions* 1–6,

(i) $\boldsymbol{B}_1 = \boldsymbol{B}_1^* + \boldsymbol{O}_P(Q^{1.5}||\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2)$, *where* $\boldsymbol{B}_1$ *and* $\boldsymbol{B}_1^*$ *are* $Q\times Q$ *matrices with* (q,q')*-th element equal to* $\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_{q'}(\bar{\boldsymbol{\rho}}))$ *and* $\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_q\boldsymbol{G}_{q'})$, *respectively*;

(ii) $\boldsymbol{B}_2 = \boldsymbol{B}_2^* + \boldsymbol{o}_P(1)$, *where* $\boldsymbol{B}_2$ *and* $\boldsymbol{B}_2^*$ *are* $Q \times Q$ *matrices with* (q, q')*-th element equal to* $\frac{1}{nT}\mathrm{tr}((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{G}_{q'}\boldsymbol{\varepsilon})$ *and* $\frac{\sigma_0^2}{n}\mathrm{tr}(\boldsymbol{G}_q'\boldsymbol{G}_{q'})$, *respectively;*

(iii)
$$\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t^* - \mathcal{H} = \frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} \; + \begin{pmatrix} \boldsymbol{B}_2^* & \boldsymbol{0}_{Q\times K} \\ \boldsymbol{0}_{K\times Q} & \boldsymbol{0}_{K\times K} \end{pmatrix}$$
$$+ \boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\right) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n, T\}}}\right),$$

where $\mathcal{H} := \frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{P}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\mathcal{Z}}^*$ *and* $\boldsymbol{B}_2^*$ *is defined in part* (ii);

(iv)
$$\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) = \frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon}) + \mathcal{H}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + \boldsymbol{\Delta}_1$$
$$+ \left(\boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n, T\}}}\right)\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0),$$

where $\boldsymbol{\Delta}_1$ *is a term of order* $\boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5}, T^{1.5}\}}\right) + \boldsymbol{O}_P\left(\frac{P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2}{\min\{\sqrt{nT}, T\}}\right) + \boldsymbol{O}_P\left(\frac{Q\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^3}{\sqrt{T}}\right)$ *and* $\mathcal{H}$ *is defined in part* (iii);

(v)
$$\hat{\sigma}^2 = (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\left(\boldsymbol{\mathcal{K}} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n, T\}}}\right)\right)(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})$$
$$+ 2(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\left(\begin{pmatrix} \boldsymbol{b}_3^* \\ \boldsymbol{0}_{K\times 1} \end{pmatrix} + \boldsymbol{O}_P\left(\sqrt{\frac{Q}{nT}}\right)\right) + \frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) + \Delta_2,$$

where $\boldsymbol{b}_3^*$ *is* $Q \times 1$ *with* q*-th element* $\frac{\sigma_0^2}{n}\mathrm{tr}(\boldsymbol{G}_q)$, $\boldsymbol{\mathcal{K}} := \frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} + \boldsymbol{B}_2^*$, $\boldsymbol{B}_2^*$ *is defined in part (ii), and* Δ_2 *has the same order as* $\boldsymbol{\Delta}_1$ in part (iv);

(vi) $\frac{1}{\sqrt{nT}}\boldsymbol{b}_4 = \frac{1}{\sqrt{nT}}\boldsymbol{b}_4^* + \boldsymbol{O}_P(\sqrt{Q})$, *where* $\boldsymbol{b}_4$ *and* $\boldsymbol{b}_4^*$ *are* $Q \times 1$ *vectors with* q*-th element equal to* $\mathrm{tr}\left((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)$ *and* $T\sigma_0^2\mathrm{tr}(\boldsymbol{G}_q)$, *respectively.*

Lemma F.4 *Under Assumptions* 1–6,

(i) $\sum_{p=1}^{P}||\sum_{k=1}^{K}(\hat{\beta}_k^0 - \beta_k)(\boldsymbol{G}_p(\hat{\boldsymbol{\rho}}) - \boldsymbol{G}_p)\boldsymbol{\mathcal{X}}_k||_F^2 = O_P(Q^2KnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^4)$;

(ii) $\sum_{p=1}^{P}||\sum_{k=1}^{K}(\hat{\beta}_k^0 - \beta_k)\boldsymbol{G}_p\boldsymbol{\mathcal{X}}_k||_F^2 = O_P(QKnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2)$;

(iii) $\sum_{p=1}^{P}||\sum_{k=1}^{K}\beta_k^0(\boldsymbol{G}_p - \boldsymbol{G}_p(\hat{\boldsymbol{\rho}}))\boldsymbol{\mathcal{X}}_k||_F^2 = O_P(Q^2KnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2)$;

(iv) $\sum_{p=1}^{P}||\sum_{k=1}^{K}(\hat{\beta}_k - \beta_k^0)\boldsymbol{G}(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k||_F^2 = O_P(QKnT||\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2^2)$;

(v) $\sum_{p=1}^{P}||\sum_{k=1}^{K}\beta_k^0\boldsymbol{G}(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k||_F^2 = O_P(QKnT)$.

Proof of Lemma A.4. Consider the first order condition (A.15)

$$\frac{\partial \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\Lambda})}{\partial \boldsymbol{\theta}} = \frac{\partial \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda})}{\partial \boldsymbol{\theta}} - \frac{\partial \varrho(\boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\zeta})}{\partial \boldsymbol{\theta}} = \mathbf{0}_{P\times 1}. \tag{F.1}$$

Evaluating (A.16) at $\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}}$,

$$\frac{\partial \mathcal{L}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} -\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1(\hat{\boldsymbol{\rho}})) \\ \vdots \\ -\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q(\hat{\boldsymbol{\rho}})) \\ \mathbf{0}_{K\times 1} \end{pmatrix} + \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{y}_t - \boldsymbol{X}_t\hat{\boldsymbol{\beta}}) =: \boldsymbol{\mathcal{P}}_1 + \boldsymbol{\mathcal{P}}_2, \tag{F.2}$$

where $\boldsymbol{Z}_t^* \coloneqq (\boldsymbol{W}_1\boldsymbol{y}_t, ..., \boldsymbol{W}_Q\boldsymbol{y}_t, \boldsymbol{X}_t)$ and $\hat{\sigma}^2 \coloneqq \hat{\sigma}^2(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}})$. First, a mean value expansion of $\boldsymbol{\mathcal{P}}_1$ around the true parameter vector $\boldsymbol{\theta}^0$ gives

$$\boldsymbol{\mathcal{P}}_1 = \begin{pmatrix} -\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1) \\ \vdots \\ -\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q) \\ \mathbf{0}_{K\times 1} \end{pmatrix} - \begin{pmatrix} \begin{pmatrix} \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})) & \cdots & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})) \\ \vdots & \ddots & \vdots \\ \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})) & \cdots & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})) \end{pmatrix} & \mathbf{0}_{Q\times K} \\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K} \end{pmatrix} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0), \tag{F.3}$$

with $\bar{\boldsymbol{\rho}} \coloneqq w\boldsymbol{\rho}^0 + (1-w)\hat{\boldsymbol{\rho}}$ for some $w \in (0,1)$. Second, substituting the true DGP into $\boldsymbol{\mathcal{P}}_2$ and expanding gives,

$$\begin{aligned} \boldsymbol{\mathcal{P}}_2 &= \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}(\boldsymbol{X}_t\boldsymbol{\beta}^0 + \boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) - \boldsymbol{X}_t\hat{\boldsymbol{\beta}}) \\ &= \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t^*(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t), \end{aligned} \tag{F.4}$$

using Lemma A.2(i). Combining (F.2), (F.3) and (F.4) gives the result

$$\begin{aligned} \frac{\partial \mathcal{L}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}})}{\partial \boldsymbol{\theta}} &= \begin{pmatrix} -\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1) \\ \vdots \\ -\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q) \\ \mathbf{0}_{K\times 1} \end{pmatrix} - \begin{pmatrix} \begin{pmatrix} \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})) & \cdots & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})) \\ \vdots & \ddots & \vdots \\ \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})) & \cdots & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})) \end{pmatrix} & \mathbf{0}_{Q\times K} \\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K} \end{pmatrix} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\ &\quad + \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t^*(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) \\ &=: \boldsymbol{\mathcal{B}}_1 - \boldsymbol{\mathcal{B}}_2(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + \frac{1}{\hat{\sigma}^2}\boldsymbol{\mathcal{B}}_3(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \frac{1}{\hat{\sigma}^2}\boldsymbol{\mathcal{B}}_4. \end{aligned} \tag{F.5}$$

Applying Lemmas F.3(i) and F.3(iv) to $\boldsymbol{\mathcal{B}}_2$ and $\boldsymbol{\mathcal{B}}_4$, respectively, and collecting terms together, the first order condition (F.1) becomes

$$\left(\begin{pmatrix} \boldsymbol{B}_1^* & \mathbf{0}_{Q\times K} \\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K} \end{pmatrix} + \frac{1}{\hat{\sigma}^2}(\boldsymbol{\mathcal{B}}_3 - \boldsymbol{\mathfrak{H}}) + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$$

$$= \frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\text{vec}(\boldsymbol{\varepsilon}) + \boldsymbol{\mathcal{B}}_1 + \boldsymbol{\Delta}_1 - \frac{\partial \varrho(\boldsymbol{\theta}, \gamma, \boldsymbol{\zeta})}{\partial \boldsymbol{\theta}}. \tag{F.6}$$

where $\boldsymbol{\mathfrak{H}}$ is defined in Lemma F.3(iii) and $\boldsymbol{B}_1^*$ is defined in Lemma F.3(i). Note, by Lemmas F.3(v) and F.1(x), $\frac{1}{\hat{\sigma}^2} = \frac{1}{\sigma_0^2} + O_P\left(\frac{\sqrt{P}}{\min\{n,T\}}\right) + O_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\right)$, and also that $||\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}||_2 \leq \frac{1}{nT}||\boldsymbol{\mathcal{Z}}||_2^2||\boldsymbol{M}_{\boldsymbol{F}^0}||_2||\boldsymbol{M}_{\boldsymbol{\Lambda}^0}||_2 = O_P(1)$ using Assumption 4.2. Using these, and applying Lemma F.3(iii) to $\boldsymbol{\mathcal{B}}_3 - \boldsymbol{\mathfrak{H}}$, (F.6) becomes

$$\begin{aligned}&\left(\frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} + \begin{pmatrix} \boldsymbol{B}_1^* + \boldsymbol{B}_2^* & \mathbf{0}_{Q\times K} \\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K} \end{pmatrix} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)\right.\\ &\left.+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\\ &= \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\text{vec}(\boldsymbol{\varepsilon}) + \boldsymbol{\mathcal{B}}_1 + \boldsymbol{\Delta}_1 - \frac{\partial \varrho(\boldsymbol{\theta}, \gamma, \boldsymbol{\zeta})}{\partial \boldsymbol{\theta}}.\end{aligned} \tag{F.7}$$

Now multiply (F.7) by $\sqrt{nT}$ to give

$$\begin{aligned}&\left(\frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} + \begin{pmatrix} \boldsymbol{B}_1^* + \boldsymbol{B}_2^* & \mathbf{0}_{Q\times K} \\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K} \end{pmatrix} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)\right.\\ &\left.+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)\sqrt{nT}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\\ &= \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\text{vec}(\boldsymbol{\varepsilon}) + \sqrt{nT}\left(\boldsymbol{\mathcal{B}}_1 + \boldsymbol{\Delta}_1 - \frac{\partial \varrho(\boldsymbol{\theta}, \gamma, \boldsymbol{\zeta})}{\partial \boldsymbol{\theta}}\right)\\ &= \frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\text{vec}(\boldsymbol{\varepsilon}) + \sqrt{nT}\left(\boldsymbol{\mathcal{B}}_1 + \boldsymbol{\Delta}_1 - \frac{\partial \varrho(\boldsymbol{\theta}, \gamma, \boldsymbol{\zeta})}{\partial \boldsymbol{\theta}}\right) + \boldsymbol{o}_P(1),\end{aligned} \tag{F.8}$$

where the last line follows by applying Lemma F.2(vii). Recalling the definition of $\boldsymbol{\mathcal{Z}}^*$, $\frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\text{vec}(\boldsymbol{\varepsilon})$ can be expanded to give

$$\begin{aligned}&\frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\text{vec}(\boldsymbol{\varepsilon})\\ &= \frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\text{vec}(\boldsymbol{\varepsilon}) + \frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\begin{pmatrix} \text{tr}\left((\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \text{tr}\left((\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \mathbf{0}_{K\times 1} \end{pmatrix}\end{aligned}$$

$$
+\frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\begin{pmatrix}\operatorname{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \operatorname{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}. \tag{F.9}
$$

Each element $\operatorname{tr}((\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0'})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})$ is zero since $\boldsymbol{M}_{\boldsymbol{F}^0}\boldsymbol{F}^0=\boldsymbol{0}_{T\times R}$. In addition, note that $\sqrt{nT}\boldsymbol{\Delta}_1=\boldsymbol{o}_P(1)$ using Assumption 6.4. Therefore, (F.8) becomes

$$
\begin{aligned}
&\left(\frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}+\begin{pmatrix}\boldsymbol{B}_1^*+\boldsymbol{B}_2^* & \boldsymbol{0}_{Q\times K}\\ \boldsymbol{0}_{K\times Q} & \boldsymbol{0}_{K\times K}\end{pmatrix}+\boldsymbol{O}_P(\sqrt{Q}P\|\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}\|_2)\right.\\
&\left.+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)\sqrt{nT}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\\
&=\frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\operatorname{vec}(\boldsymbol{\varepsilon})+\frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\begin{pmatrix}\operatorname{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \operatorname{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\\
&\quad-\begin{pmatrix}\sqrt{\frac{T}{n}}\operatorname{tr}(\boldsymbol{G}_1)\\ \vdots \\ \sqrt{\frac{T}{n}}\operatorname{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}-\sqrt{nT}\frac{\partial\varrho(\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\zeta})}{\partial\boldsymbol{\theta}}+\boldsymbol{o}_P(1)\\
&=\frac{1}{\hat{\sigma}^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\operatorname{vec}(\boldsymbol{\varepsilon})\\
&\quad+\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\left(\begin{pmatrix}\operatorname{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \operatorname{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}-\hat{\sigma}^2\begin{pmatrix}T\operatorname{tr}(\boldsymbol{G}_1)\\ \vdots \\ T\operatorname{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\right)-\sqrt{nT}\frac{\partial\varrho(\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\zeta})}{\partial\boldsymbol{\theta}}+\boldsymbol{o}_P(1).\\
&\quad+\left(\frac{1}{\hat{\sigma}^2}-\frac{1}{\sigma_0^2}\right)\frac{1}{\sqrt{nT}}\left(\begin{pmatrix}\operatorname{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \operatorname{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}-\hat{\sigma}^2\begin{pmatrix}T\operatorname{tr}(\boldsymbol{G}_1)\\ \vdots \\ T\operatorname{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\right),
\end{aligned} \tag{F.10}
$$

where the second equality follows by adding and subtracting terms. Using Lemmas F.3(v), F.3(vi) and F.1(x), the last term in (F.10) $\boldsymbol{o}_P(1)$. In addition, using Lemmas F.3(v) and

F.3(vi) again,

$$\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\left(\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix} - \hat{\sigma}^2\begin{pmatrix}T\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ T\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\right)$$

$$= \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\left(\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix} - \mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\begin{pmatrix}\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\right)$$

$$- \frac{2}{\sigma_0^2}\frac{1}{\sqrt{nT}}(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\left(\begin{pmatrix}\boldsymbol{b}_3^*\\ \boldsymbol{0}_{K\times 1}\end{pmatrix} + \boldsymbol{O}_P\left(\sqrt{\frac{Q}{nT}}\right)\right)\begin{pmatrix}T\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ T\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}$$

$$- \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\left(\boldsymbol{\mathcal{K}} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\begin{pmatrix}T\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ T\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}$$

$$- \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\Delta_2\begin{pmatrix}T\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ T\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}$$

$$= \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\left(\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix} - \mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\begin{pmatrix}\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\right)$$

$$- \frac{2}{\sigma_0^2}\frac{1}{\sqrt{nT}}\begin{pmatrix}T\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ T\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\begin{pmatrix}\frac{\sigma_0^2}{n}\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots \\ \frac{\sigma_0^2}{n}\mathrm{tr}(\boldsymbol{G}_Q)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}'(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \boldsymbol{o}_P(1)$$

$$+ \Bigg(\boldsymbol{O}_p(Q) + \boldsymbol{O}_P(\sqrt{Q}\sqrt{nT}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2 + \boldsymbol{O}_P(QP\sqrt{nT}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2)$$

$$+\boldsymbol{O}_P\left(\frac{\sqrt{Q}P\sqrt{nT}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right)\Bigg)(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}). \tag{F.11}$$

Using (F.11) in (F.10), and ignoring dominated terms, gives the result,

$$\Bigg(\frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}+\begin{pmatrix}\boldsymbol{B}_1^*+\boldsymbol{B}_2^* & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix}-\frac{2}{n^2}\begin{pmatrix}\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots\\ \mathrm{tr}(\boldsymbol{G}_Q)\\ \mathbf{0}_{K\times 1}\end{pmatrix}\begin{pmatrix}\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots\\ \mathrm{tr}(\boldsymbol{G}_Q)\\ \mathbf{0}_{K\times 1}\end{pmatrix}'$$

$$+\boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\Bigg)\sqrt{nT}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)$$

$$=\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon})$$

$$+\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\left(\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots\\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \mathbf{0}_{K\times 1}\end{pmatrix}-\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\begin{pmatrix}\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1)\\ \vdots\\ \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q)\\ \mathbf{0}_{K\times 1}\end{pmatrix}\right)$$

$$-\sqrt{nT}\frac{\partial\varrho(\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\zeta})}{\partial\boldsymbol{\theta}}+\boldsymbol{o}_P(1)$$

$$=\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon})+\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots\\ \mathrm{tr}\left((\boldsymbol{G}_Q^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \mathbf{0}_{K\times 1}\end{pmatrix}$$

$$-\sqrt{nT}\frac{\partial\varrho(\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\zeta})}{\partial\boldsymbol{\theta}}+\boldsymbol{o}_P(1), \tag{F.12}$$

where the last equality follows by recalling the definition $\boldsymbol{G}_q^* := \boldsymbol{G}_q - \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_q)\boldsymbol{I}_n$. For the second to last term note,

$$\begin{aligned}\left|\left|\sqrt{nT}\frac{\partial \varrho(\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\zeta})}{\partial \boldsymbol{\theta}}\right|\right|_2 &= \sqrt{nT}\left(\sum_{p=1}^{P}\left(\gamma_p \frac{1}{|\theta_p^\dagger|^{\zeta_p}}\frac{\hat{\theta}_p}{|\hat{\theta}_p|}\right)^2\right)^{\frac{1}{2}} \\ &\leq \sqrt{PnT}\left|\frac{\gamma_p}{|\theta_p^\dagger|^{\zeta_p}}\right| = o(1), \end{aligned} \tag{F.13}$$

where the inequality in (F.13) follows the same steps as those used to obtain (D.26), and the final line follows under Assumption 6.6. Moreover, recalling the definition given in (16), notice that

$$\frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} + \begin{pmatrix} \boldsymbol{B}_1^* + \boldsymbol{B}_2^* & \boldsymbol{0}_{Q\times K} \\ \boldsymbol{0}_{K\times Q} & \boldsymbol{0}_{K\times K}\end{pmatrix} - \frac{2}{n^2}\begin{pmatrix}\mathrm{tr}(\boldsymbol{G}_1) \\ \vdots \\ \mathrm{tr}(\boldsymbol{G}_Q) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\begin{pmatrix}\mathrm{tr}(\boldsymbol{G}_1) \\ \vdots \\ \mathrm{tr}(\boldsymbol{G}_Q) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix}' = \boldsymbol{D}. \tag{F.14}$$

Hence, applying (F.13) and (F.14) to (F.12) gives the result

$$\begin{aligned}&\left(\boldsymbol{D} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)\sqrt{nT}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\ &= \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) + \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} + \boldsymbol{o}_P(1). \end{aligned} \tag{F.15}$$

By Lemma F.2(viii), $\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) = O_P(\sqrt{P})$, and by Lemma F.3(vi)

$$\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} = \sqrt{\frac{T}{n}}\begin{pmatrix}\mathrm{tr}\left(\boldsymbol{G}_1^*\right) \\ \vdots \\ \mathrm{tr}\left(\boldsymbol{G}_Q^*\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} = \boldsymbol{0}_{P\times 1} + \boldsymbol{O}_P(\sqrt{Q}) \tag{F.16}$$

because $\mathrm{tr}(\boldsymbol{G}_q^*) = 0$ for $q = 1, ..., Q$. Thus,

$$\left(\boldsymbol{D} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)\sqrt{nT}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$$

$$= \boldsymbol{O}_P(\sqrt{P}) + \boldsymbol{O}_P(\sqrt{Q}) + \boldsymbol{o}_P(1). \tag{F.17}$$

Since $||\boldsymbol{D}||_2, ||\boldsymbol{D}^{-1}||_2 = O_P(1)$, and by Proposition 1, $||\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2$ is at least of order a_{nT} then $||\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2 = O_P\left(\sqrt{\frac{P}{nT}}\right)$ follows using Assumption 6.1, and the final result is obtained,

$$\begin{aligned}\boldsymbol{D}\sqrt{nT}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) &= \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\text{vec}(\boldsymbol{\varepsilon}) \\ &+ \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\begin{pmatrix}\text{tr}\left((\boldsymbol{G}_1^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \text{tr}\left((\boldsymbol{G}_Q^*\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} + \boldsymbol{o}_P(1). \end{aligned}\tag{F.18}$$

□

Proof of Lemma A.5(i). From the definition of $\boldsymbol{D}$ given in equation (16) of the main text,

$$\begin{aligned}\boldsymbol{D} - \hat{\boldsymbol{D}} = &\frac{1}{\hat{\sigma}^2}\frac{1}{nT}(\boldsymbol{\mathcal{Z}} - \hat{\boldsymbol{\mathcal{Z}}})'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} + \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\hat{\boldsymbol{\mathcal{Z}}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})(\boldsymbol{\mathcal{Z}} - \hat{\boldsymbol{\mathcal{Z}}}) \\ &+ \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\hat{\boldsymbol{\mathcal{Z}}}'((\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0}) \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\hat{\boldsymbol{\mathcal{Z}}} + \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\hat{\boldsymbol{\mathcal{Z}}}'(\boldsymbol{M}_{\hat{\boldsymbol{F}}} \otimes (\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}))\hat{\boldsymbol{\mathcal{Z}}} \\ &+ \left(\frac{1}{\sigma_0^2} - \frac{1}{\hat{\sigma}^2}\right)\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} + \begin{pmatrix}\boldsymbol{\Omega} - \hat{\boldsymbol{\Omega}} & \boldsymbol{0}_{Q\times K} \\ \boldsymbol{0}_{K\times Q} & \boldsymbol{0}_{Q\times Q}\end{pmatrix} \\ =: &\boldsymbol{L}_1 + ... + \boldsymbol{L}_6. \end{aligned}\tag{F.19}$$

Note that, for $p = 1, ..., Q$,

$$\begin{aligned}\boldsymbol{\mathfrak{Z}}_p - \hat{\boldsymbol{\mathfrak{Z}}}_p &= \sum_{k=1}^{K}(\hat{\beta}_k^0 - \beta_k)(\boldsymbol{G}_p(\hat{\boldsymbol{\rho}}) - \boldsymbol{G}_p)\boldsymbol{\mathcal{X}}_k + \sum_{k=1}^{K}(\hat{\beta}_k^0 - \beta_k)\boldsymbol{G}_p\boldsymbol{\mathcal{X}}_k + \sum_{k=1}^{K}\beta_k^0(\boldsymbol{G}_p - \boldsymbol{G}_p(\hat{\boldsymbol{\rho}}))\boldsymbol{\mathcal{X}}_k \\ &=: (\boldsymbol{\mathfrak{Z}}_p - \hat{\boldsymbol{\mathfrak{Z}}}_p)^{(1)} + (\boldsymbol{\mathfrak{Z}}_p - \hat{\boldsymbol{\mathfrak{Z}}}_p)^{(2)} + (\boldsymbol{\mathfrak{Z}}_p - \hat{\boldsymbol{\mathfrak{Z}}}_p)^{(3)} \end{aligned}\tag{F.20}$$

and $\boldsymbol{\mathfrak{Z}}_p - \hat{\boldsymbol{\mathfrak{Z}}}_p = \boldsymbol{0}_{n\times T}$ otherwise, and also that, for $p = 1, ..., Q$,

$$\hat{\boldsymbol{\mathfrak{Z}}}_p = \sum_{k=1}^{K}\hat{\beta}_k\boldsymbol{G}(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k = \sum_{k=1}^{K}(\hat{\beta}_k - \beta_k^0)\boldsymbol{G}(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k + \sum_{k=1}^{K}\beta_k^0\boldsymbol{G}(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k =: \hat{\boldsymbol{\mathfrak{Z}}}_p^{(1)} + \hat{\boldsymbol{\mathfrak{Z}}}_p^{(2)} \tag{F.21}$$

and $\hat{\mathcal{Z}}_p = \mathcal{Z}_p$ otherwise. As such, using (F.20) and (F.21), terms $\boldsymbol{L}_1, ..., \boldsymbol{L}_6$ in (F.19) can be expanded. For term $\boldsymbol{L}_1$,

$$\begin{aligned}
\boldsymbol{L}_1 = &\frac{1}{\hat{\sigma}^2}\frac{1}{nT}\begin{pmatrix} \mathrm{tr}\left((\mathcal{Z}_1 - \hat{\mathcal{Z}}_1)^{(1)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_1'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \mathrm{tr}\left((\mathcal{Z}_1 - \hat{\mathcal{Z}}_1)^{(1)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_P'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \\ \vdots & \ddots & \vdots \\ \mathrm{tr}\left((\mathcal{Z}_P - \hat{\mathcal{Z}}_P)^{(1)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_1'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \mathrm{tr}\left((\mathcal{Z}_P - \hat{\mathcal{Z}}_P)^{(1)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_P'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \end{pmatrix} \\
&+ \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\begin{pmatrix} \mathrm{tr}\left((\mathcal{Z}_1 - \hat{\mathcal{Z}}_1)^{(2)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_1'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \mathrm{tr}\left((\mathcal{Z}_1 - \hat{\mathcal{Z}}_1)^{(2)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_P'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \\ \vdots & \ddots & \vdots \\ \mathrm{tr}\left((\mathcal{Z}_P - \hat{\mathcal{Z}}_P)^{(2)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_1'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \mathrm{tr}\left((\mathcal{Z}_P - \hat{\mathcal{Z}}_P)^{(2)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_P'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \end{pmatrix} \\
&+ \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\begin{pmatrix} \mathrm{tr}\left((\mathcal{Z}_1 - \hat{\mathcal{Z}}_1)^{(3)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_1'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \mathrm{tr}\left((\mathcal{Z}_1 - \hat{\mathcal{Z}}_1)^{(3)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_P'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \\ \vdots & \ddots & \vdots \\ \mathrm{tr}\left((\mathcal{Z}_P - \hat{\mathcal{Z}}_P)^{(3)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_1'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \mathrm{tr}\left((\mathcal{Z}_P - \hat{\mathcal{Z}}_P)^{(3)}\boldsymbol{M}_{\boldsymbol{F}^0}\mathcal{Z}_P'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \end{pmatrix} \\
=: &\frac{1}{\hat{\sigma}^2}\frac{1}{nT}(\boldsymbol{L}_{1.1} + \boldsymbol{L}_{1.2} + \boldsymbol{L}_{1.3}). \qquad \text{(F.22)}
\end{aligned}$$

Using Lemmas F.4(i),...,F.4(iii), and the inequalities $\mathrm{tr}(\boldsymbol{AB}) \le ||\boldsymbol{A}||_F||\boldsymbol{B}||_F$ and $||\boldsymbol{AB}||_F \le ||\boldsymbol{A}||_2||\boldsymbol{B}||_F$,

$$\begin{aligned}
||\boldsymbol{L}_{1.1}||_F &\le \left(\sum_{p=1}^{P}||(\mathcal{Z}_p - \hat{\mathcal{Z}}_p)^{(1)}||_F^2\right)^{\frac{1}{2}}\left(\sum_{p'=1}^{P}||\mathcal{Z}_{p'}||_F^2\right)^{\frac{1}{2}} \\
&= O_P(Q\sqrt{KP}nT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2), \qquad \text{(F.23)}
\end{aligned}$$

$$\begin{aligned}
||\boldsymbol{L}_{1.2}||_F &\le \left(\sum_{p=1}^{P}||(\mathcal{Z}_p - \hat{\mathcal{Z}}_p)^{(2)}||_F^2\right)^{\frac{1}{2}}\left(\sum_{p'=1}^{P}||\mathcal{Z}_{p'}||_F^2\right)^{\frac{1}{2}} \\
&= O_P(\sqrt{QKP}nT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2), \qquad \text{(F.24)}
\end{aligned}$$

$$\begin{aligned}
||\boldsymbol{L}_{1.3}||_F &\le \left(\sum_{p=1}^{P}||(\mathcal{Z}_p - \hat{\mathcal{Z}}_p)^{(3)}||_F^2\right)^{\frac{1}{2}}\left(\sum_{p'=1}^{P}||\mathcal{Z}_{p'}||_F^2\right)^{\frac{1}{2}} \\
&= O_P(Q\sqrt{KP}nT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2). \qquad \text{(F.25)}
\end{aligned}$$

Thus, $||\boldsymbol{L}_1||_F = O_P(Q\sqrt{KP}nT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$. $||\boldsymbol{L}_2||_F$ is of the same order. For term $\boldsymbol{L}_3$, this can be expanded as

$$\begin{aligned}
\boldsymbol{L}_3 =: &\frac{1}{\hat{\sigma}^2}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \\ \vdots & \ddots & \vdots \\ \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \end{pmatrix} \\
&+ \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \\ \vdots & \ddots & \vdots \\ \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \end{pmatrix} \\
&+ \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \\ \vdots & \ddots & \vdots \\ \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(1)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \end{pmatrix} \\
&+ \frac{1}{\hat{\sigma}^2}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \\ \vdots & \ddots & \vdots \\ \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_1^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) & \cdots & \operatorname{tr}\left(\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)}(\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0})\hat{\boldsymbol{\mathfrak{Z}}}_P^{(2)'}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \end{pmatrix} \\
=: &\frac{1}{\hat{\sigma}^2}\frac{1}{nT}(\boldsymbol{L}_{3.1} + \boldsymbol{L}_{3.2} + \boldsymbol{L}_{3.3} + \boldsymbol{L}_{3.4}). \qquad \text{(F.26)}
\end{aligned}$$

Considering each of the four terms in (F.26),

$$\begin{aligned}
||\boldsymbol{L}_{3.1}||_F &\leq ||\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0}||_2 \left(\sum_{p=1}^{P} ||\hat{\boldsymbol{\mathfrak{Z}}}_p^{(1)}||_F^2\right) \\
&= O_P(Q^{1.5}KnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^3) + O_P\left(\frac{QKnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2}{\sqrt{\min\{n, T\}}}\right), \qquad \text{(F.27)}
\end{aligned}$$

$$\begin{aligned}
||\boldsymbol{L}_{3.2}||_F &\leq ||\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0}||_2 \left(\sum_{p=1}^{P} ||\hat{\boldsymbol{\mathfrak{Z}}}_p^{(1)}||_F^2\right)^{\frac{1}{2}} \left(\sum_{p'=1}^{P} ||\hat{\boldsymbol{\mathfrak{Z}}}_{p'}^{(2)}||_F^2\right)^{\frac{1}{2}} \\
&= O_P\left(Q^{1.5}KnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2\right) + O_P\left(\frac{QKnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n, T\}}}\right), \qquad \text{(F.28)}
\end{aligned}$$

$||\boldsymbol{L}_{3.3}||_F$ has the same order as $||\boldsymbol{L}_{3.2}||_F$, and

$$||\boldsymbol{L}_{3.4}||_F \leq ||\boldsymbol{P}_{\hat{\boldsymbol{F}}} - \boldsymbol{P}_{\boldsymbol{F}^0}||_2 \left(\sum_{p=1}^{P} ||\hat{\boldsymbol{\mathfrak{Z}}}_p^{(2)}||_F^2\right)$$

$$= O_P(Q^{1.5}KnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{QKnT}{\sqrt{\min\{n,T\}}}\right), \tag{F.29}$$

where the above uses Lemmas F.2(v), F.4(iv) and F.4(v). Thus $||\boldsymbol{L}_3||_F = O_P(Q^{1.5}KnT||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{QKnT}{\sqrt{\min\{n,T\}}}\right)$. It can be shown that $||\boldsymbol{L}_4||_F$ is of the same order. For term $\boldsymbol{L}_5$, $||\boldsymbol{L}_5||_2 \leq |\sigma_0^{-2} - \hat{\sigma}^{-2}|||\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}||_2$. By combining Lemmas F.3(v) and F.1(x), $|\sigma_0^{-2} - \hat{\sigma}^{-2}| = \frac{1}{\sigma_0^2} + O_P\left(\frac{\sqrt{P}}{\min\{n,T\}}\right) + O_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\right)$, and, by Assumption 4.2, $||\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}||_2 = O_P(1)$. Finally, let $\boldsymbol{\Omega} - \hat{\boldsymbol{\Omega}} =: \hat{\boldsymbol{\Omega}}^{(1)} + ... + \hat{\boldsymbol{\Omega}}^{(6)}$ with elements $(\boldsymbol{\Omega}^{(1)})_{qq'} = \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_q(\boldsymbol{G}_{q'} - \boldsymbol{G}_{q'}(\hat{\boldsymbol{\rho}}))$, $(\boldsymbol{\Omega}^{(2)})_{qq'} = \frac{1}{n}\mathrm{tr}((\boldsymbol{G}_q - \boldsymbol{G}_q(\hat{\boldsymbol{\rho}}))\boldsymbol{G}_{q'}(\hat{\boldsymbol{\rho}}))$, $(\boldsymbol{\Omega}^{(3)})_{qq'} = \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_q(\boldsymbol{G}'_{q'} - \boldsymbol{G}'_{q'}(\hat{\boldsymbol{\rho}})))$, $(\boldsymbol{\Omega}^{(4)})_{qq'} = \frac{1}{n}\mathrm{tr}((\boldsymbol{G}_q - \boldsymbol{G}_q(\hat{\boldsymbol{\rho}}))\boldsymbol{G}'_{q'}(\hat{\boldsymbol{\rho}}))$, $(\boldsymbol{\Omega}^{(5)})_{qq'} = \frac{2}{n^2}\mathrm{tr}(\boldsymbol{G}_q(\hat{\boldsymbol{\rho}}) - \boldsymbol{G}_q)\mathrm{tr}(\boldsymbol{G}_{q'}(\hat{\boldsymbol{\rho}}))$ and $(\boldsymbol{\Omega}^{(6)})_{qq'} = \frac{2}{n^2}\mathrm{tr}(\boldsymbol{G}_q)\mathrm{tr}(\boldsymbol{G}_{q'}(\hat{\boldsymbol{\rho}}) - \boldsymbol{G}_{q'})$. For the first of these terms,

$$\begin{aligned}
||\hat{\boldsymbol{\Omega}}^{(1)}||_F &\leq ||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n||_2 \left(\sum_{q=1}^{Q}||\boldsymbol{G}_q||_2^2\right)^{\frac{1}{2}} \left(\sum_{q'=1}^{Q}||\boldsymbol{G}_{q'}(\hat{\boldsymbol{\rho}})||_2\right)^{\frac{1}{2}} \\
&= O_P(Q^{1.5}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2), \qquad \text{(F.30)}
\end{aligned}$$

using Lemma A.2(viii), and so $||\hat{\boldsymbol{\Omega}}^{(1)}||_F = O_P(Q^{1.5}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$. Similar steps establish that $||\hat{\boldsymbol{\Omega}}^{(2)}||_F^2, ..., ||\hat{\boldsymbol{\Omega}}^{(5)}||_F^2$ have the same order, and therefore $||\boldsymbol{L}_6||_F = O_P(Q^{1.5}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$. Combining the above results for $||\boldsymbol{L}_1||_F, ..., ||\boldsymbol{L}_6||_F$ yields $||\boldsymbol{D}^{-1} - \hat{\boldsymbol{D}}^{-1}||_2 = O_P(Q^{1.5}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{QP}{\sqrt{\min\{n,T\}}}\right)$. □

Proof of Lemma A.5(ii). Notice, firstly, that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{q=1}^{Q}\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}) - \sigma_0^2 T\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right] \\
=&\sum_{q=1}^{Q}\mathbb{E}\left[\mathbb{E}_{\mathcal{D}}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}) - \sigma_0^2 T\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right]\right]. \qquad \text{(F.31)}
\end{aligned}$$

Next, since $\mathbb{E}_{\mathcal{D}}\left[\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})\right] = \mathbb{E}_{\mathcal{D}}\left[\mathrm{vec}(\boldsymbol{\varepsilon})'(\boldsymbol{I}_T \otimes \boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\mathrm{vec}(\boldsymbol{\varepsilon})\right] = \sigma_0^2 T\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)$, and, using the same steps as in Lemma 3 of Yu et al. (2008),

$$\begin{aligned}
&\mathbb{E}_{\mathcal{D}}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})\right)^2\right] \\
=&(\mathcal{M}_\varepsilon^4 - 3\sigma_0^4)T\sum_{i=1}^{n}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)_{ii}^2 + \sigma_0^4(\mathrm{tr}(\boldsymbol{I}_T \otimes \boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*))^2
\end{aligned}$$

$$+\sigma_0^4\mathrm{tr}((\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)')+\sigma_0^4\mathrm{tr}((\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)), \quad (F.32)$$

then

$$\begin{aligned}&\sum_{q=1}^{Q}\mathbb{E}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})-\sigma_0^2T\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right]\\ =&\sum_{q=1}^{Q}\mathbb{E}\Big[(\mathcal{M}_\varepsilon^4-3\sigma_0^4)T\sum_{i=1}^{n}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)_{ii}^2+\sigma_0^4\mathrm{tr}((\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)')\\ &+\sigma_0^4\mathrm{tr}((\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*))\Big]. \quad (F.33)\end{aligned}$$

Now note that $\mathbb{E}\left[\mathrm{tr}((\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)')\right]=\mathbb{E}\left[\mathrm{tr}(\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)')\right]=T\mathbb{E}\left[\mathrm{tr}((\boldsymbol{G}_q^*)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right]=T\mathbb{E}\left[||(\boldsymbol{G}_q^*)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)||_F^2\right]\leq T||\boldsymbol{G}_q^*||_2^4\mathbb{E}\left[||\boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_F^2\right]\leq TR^0||\boldsymbol{G}_q^*||_2^4=O(T)$ since $\boldsymbol{G}_q^*$ is UB over q. The same also applies to $\mathbb{E}\left[\mathrm{tr}((\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{I}_T\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*))\right]$, and similarly for the first term $\mathbb{E}\left[\sum_{i=1}^n(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)_{ii}^2\right]$ as $\sum_{i=1}^n(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)_{ii}^2\leq||\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*||_F^2$. Hence

$$\mathbb{E}\left[\sum_{q=1}^{Q}\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})-\sigma_0^2T\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right]=O(QT). \quad (F.34)$$

□

Proof of Lemma A.5(iii). Notice, firstly, that

$$\begin{aligned}&\mathbb{E}\left[\sum_{q=1}^{Q}\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})-\sigma_0^2R^0\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right]\\ =&\sum_{q=1}^{Q}\mathbb{E}\left[\mathbb{E}_{\mathcal{D}}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})-\sigma_0^2R^0\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right]\right]. \quad (F.35)\end{aligned}$$

Next, since $\mathbb{E}_{\mathcal{D}}\left[\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\right]=\mathbb{E}_{\mathcal{D}}\left[\mathrm{vec}(\boldsymbol{\varepsilon})'(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\mathrm{vec}(\boldsymbol{\varepsilon})\right]=\sigma_0^2\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)=\sigma_0^2R^0\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)$, and, using the same steps as in Lemma 3 of Yu et al. (2008),

$$\begin{aligned}&\mathbb{E}_{\mathcal{D}}\left[\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})^2\right]\\ =&(\mathcal{M}_\varepsilon^4-3\sigma_0^4)\sum_{i=1}^{nT}(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)_{ii}^2+\sigma_0^4(R^0)^2(\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*))^2\\ &+\sigma_0^4\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*))+\sigma_0^4\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)'), \quad (F.36)\end{aligned}$$

then

$$\sum_{q=1}^{Q}\mathbb{E}\left[\mathbb{E}_{\mathcal{D}}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})-\sigma_0^2R^0\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right]\right]$$

$$=\sum_{q=1}^{Q}\mathbb{E}\Big[(\mathcal{M}_\varepsilon^4-3\sigma_0^4)\sum_{i=1}^{nT}(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)_{ii}^2+\sigma_0^4\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*))$$

$$+\sigma_0^4\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)')\Big]. \tag{F.37}$$

Now note that $\mathbb{E}[\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)')]=\mathbb{E}[\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)]=R^0\mathbb{E}[\mathrm{tr}((\boldsymbol{G}_q^*)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)]=R\mathbb{E}[||(\boldsymbol{G}_q^*)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*||_F^2]\leq R^0||\boldsymbol{G}_q^*||_2^4\mathbb{E}||\boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_F^2]\leq(R^0)^2||\boldsymbol{G}_q^*||_2^4=O(1)$ because $\boldsymbol{G}_q^*$ is UB over q. Similarly for the remaining terms in (F.37). Hence

$$\mathbb{E}\left[\sum_{q=1}^{Q}\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})-\sigma_0^2R^0\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*)\right)^2\right]=O(Q). \tag{F.38}$$

□

Proof of Lemma A.5(iv). Notice, firstly, that

$$\mathbb{E}\left[\sum_{q=1}^{Q}\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})-\sigma_0^2R^0\mathrm{tr}(\boldsymbol{G}_q^*)\right)^2\right]$$

$$=\sum_{q=1}^{Q}\mathbb{E}\left[\mathbb{E}_{\mathcal{D}}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})-\sigma_0^2R^0\mathrm{tr}(\boldsymbol{G}_q^*)\right)^2\right]\right]. \tag{F.39}$$

Next, since $\mathbb{E}_{\mathcal{D}}\left[\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\right]=\mathbb{E}_{\mathcal{D}}\left[\mathrm{vec}(\boldsymbol{\varepsilon})'(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)')\mathrm{vec}(\boldsymbol{\varepsilon})\right]=\sigma_0^2\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)')=\sigma_0^2R^0\mathrm{tr}(\boldsymbol{G}_q^*)$, and, using the same steps as in Lemma 3 of Yu et al. (2008),

$$\mathbb{E}_{\mathcal{D}}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})\right)^2\right]$$

$$=(\mathcal{M}_\varepsilon^4-3\sigma_0^4)\sum_{i=1}^{nT}(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)')_{ii}^2+\sigma_0^4T^2(\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_q^*))^2$$

$$+\sigma_0^4\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)')(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)')')+\sigma_0^4\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)')(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes(\boldsymbol{G}_q^*)')), \tag{F.40}$$

then

$$\sum_{q=1}^{Q}\mathbb{E}\left[\mathbb{E}_{\mathcal{D}}\left[\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})-\sigma_0^2R^0\mathrm{tr}(\boldsymbol{G}_q^*)\right)^2\right]\right]$$

$$
\begin{aligned}
&= \sum_{q=1}^{Q} \mathbb{E}\Big[(\mathcal{M}_{\varepsilon}^4 - 3\sigma_0^4)\sum_{i=1}^{nT}(\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)')_{ii}^2 + \sigma_0^4 \mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)')(\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)')') \\
&\quad + \sigma_0^4 \mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)')(\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)'))\Big]. \qquad \text{(F.41)}
\end{aligned}
$$

Now note that $\mathbb{E}[\mathrm{tr}((\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)')(\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)')')] = \mathbb{E}[\mathrm{tr}(\boldsymbol{P}_{\boldsymbol{F}^0} \otimes (\boldsymbol{G}_q^*)'\boldsymbol{G}_q^*)] = R^0 \mathrm{tr}((\boldsymbol{G}_q^*)'\boldsymbol{G}_q^*) = R^0 ||\boldsymbol{G}_q^*||_F^2 \leq R^0 n ||\boldsymbol{G}_q^*||_2^2 = O(n)$ because $\boldsymbol{G}_q^*$ is UB over q. Similarly for the remaining terms in (F.41). Hence

$$
\mathbb{E}\left[\sum_{q=1}^{Q}\left(\mathrm{tr}((\boldsymbol{G}_q^*\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) - \sigma_0^2 R^0 \mathrm{tr}(\boldsymbol{G}_q^*)\right)^2\right] = O(Qn). \qquad \text{(F.42)}
$$

□

Proof of Lemma A.5(v). Expanding,

$$
\begin{aligned}
&\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\begin{pmatrix} \mathrm{tr}\left((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon} + \boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\right) \\ \vdots \\ \mathrm{tr}\left((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon} + \boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\right) \end{pmatrix} \\
=&\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\left(\begin{pmatrix} \mathrm{tr}\left((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\right) \\ \vdots \\ \mathrm{tr}\left((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\right) \end{pmatrix} - \begin{pmatrix} \mathrm{tr}\left((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right) \end{pmatrix} + \begin{pmatrix} \mathrm{tr}\left((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right) \end{pmatrix}\right) \\
=:&\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}(\boldsymbol{l}_1 - \boldsymbol{l}_2 + \boldsymbol{l}_3). \qquad \text{(F.43)}
\end{aligned}
$$

First,

$$
\begin{aligned}
||\boldsymbol{l}_1||_2^2 &= \frac{1}{n^2}\sum_{p=1}^{P}\mathrm{tr}\left(\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}(\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'\boldsymbol{\Lambda}^0\right)^2 \\
&\leq \frac{(R^0)^2}{n^2}\left|\left|\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right|\right|_2^2 \sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}(\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'\boldsymbol{\Lambda}^0||_2^2 \qquad \text{(F.44)}
\end{aligned}
$$

Now notice,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}(\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'\boldsymbol{\Lambda}^0||_2^2\right] \\
&\leq \sum_{p=1}^{P}\mathbb{E}\left[||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}(\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'\boldsymbol{\Lambda}^0||_F^2\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{p=1}^{P} \mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{j=1}^{n}\lambda_{ir}^0\varepsilon_{it}(\mathcal{Z}_p-\bar{\mathcal{Z}}_p)_{jt}\lambda_{js}^0\right)^2\right] \\
&= \sum_{p=1}^{P} \mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{j=1}^{n}\sum_{i'=1}^{n}\sum_{t'=1}^{T}\sum_{j'=1}^{n}\lambda_{ir}^0\lambda_{js}^0\lambda_{i'r}^0\lambda_{j's}^0\mathbb{E}_{\mathcal{D}}\left[\varepsilon_{it}(\mathcal{Z}_p-\bar{\mathcal{Z}}_p)_{jt}\varepsilon_{i't'}(\mathcal{Z}_p-\bar{\mathcal{Z}}_p)_{j't'}\right]\right].
\end{aligned} \tag{F.45}
$$

Consider, for example, the case where $p=1$. In this case $\mathcal{Z}_1-\bar{\mathcal{Z}}_1=\boldsymbol{W}_1\sum_{h=1}^{\infty}\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_h^*$. It is straightforward to see that $\mathbb{E}_{\mathcal{D}}\left[\varepsilon_{it}(\mathcal{Z}_p-\bar{\mathcal{Z}}_p)_{jt}\varepsilon_{i't'}(\mathcal{Z}_p-\bar{\mathcal{Z}}_p)_{j't'}\right]\neq 0$ only when $t=t', i=i'$ and $j=j'$, hence,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}(\mathcal{Z}_p-\bar{\mathcal{Z}}_p)'\boldsymbol{\Lambda}^0||_2^2\right] \\
&\leq \sigma_0^4\sum_{p=1}^{P}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{j=1}^{n}\lambda_{ir}^{0^2}\lambda_{js}^{0^2}\sum_{h=1}^{\infty}\sum_{i=1}^{n}(\boldsymbol{W}_1\boldsymbol{A}^h\boldsymbol{S}^{-1})_{jl}\right]=O(n^2T)
\end{aligned} \tag{F.46}
$$

using Assumptions 1, 2.2, and 2.3. Similarly for other p, trivially so in the case where p corresponds to an exogenous covariate since $\mathcal{Z}_p-\bar{\mathcal{Z}}_p=\boldsymbol{0}_{n\times T}$. Using this and (F.44) gives the result that $||\boldsymbol{l}_1||_2^2=O_P(PT)$ whereby $\frac{1}{\sqrt{nT}}||\boldsymbol{l}_1||_2=O_P(\sqrt{P/n})=o_P(1)$. For term $\boldsymbol{l}_2$,

$$
\begin{aligned}
||\boldsymbol{l}_2||_2^2 &= \sum_{p=1}^{P}\operatorname{tr}\left((\mathcal{Z}_p-\bar{\mathcal{Z}}_p)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)^2 \\
&\leq (R^0)^2\sum_{p=1}^{P}||\mathcal{Z}_p-\bar{\mathcal{Z}}_p||_2^2||\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}||_2^2 \\
&\leq \frac{(R^0)^2}{n^2T^2}\left(\sum_{p=1}^{P}||\mathcal{Z}_p-\bar{\mathcal{Z}}_p||_2^2\right)||\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{F}^0||_2^2\left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2^2\left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2^2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2^2 \\
&= O_P(P\max\{n,T\}),
\end{aligned} \tag{F.47}
$$

using Lemmas A.2(iii), F.1(ii) and F.1(v), as well as Assumptions 6.2 and 6.3. This implies $\frac{1}{\sqrt{nT}}||\boldsymbol{l}_2||_2=O_P\left(\sqrt{\frac{P}{\min\{n,T\}}}\right)=o_P(1)$. Finally for term $\boldsymbol{l}_3$,

$$
\frac{1}{\sqrt{nT}}\boldsymbol{l}_3=\frac{1}{\sqrt{nT}}\begin{pmatrix}\operatorname{tr}\left((\mathcal{Z}_1-\bar{\mathcal{Z}}_1)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)-\mathbb{E}_{\mathcal{D}}[\operatorname{tr}\left((\mathcal{Z}_1-\bar{\mathcal{Z}}_1)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)] \\ \vdots \\ \operatorname{tr}\left((\mathcal{Z}_P-\bar{\mathcal{Z}}_P)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)-\mathbb{E}_{\mathcal{D}}[\operatorname{tr}\left((\mathcal{Z}_P-\bar{\mathcal{Z}}_P)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)]\end{pmatrix}
$$

$$+\frac{1}{\sqrt{nT}}\begin{pmatrix}\mathbb{E}_{\mathcal{D}}[\mathrm{tr}\left((\mathfrak{Z}_1-\bar{\mathfrak{Z}}_1)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)]\\ \vdots \\ \mathbb{E}_{\mathcal{D}}[\mathrm{tr}\left((\mathfrak{Z}_P-\bar{\mathfrak{Z}}_P)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)]\end{pmatrix}$$

$$:=\frac{1}{\sqrt{nT}}\boldsymbol{l}_{3.1}+\frac{1}{\sqrt{nT}}\begin{pmatrix}\mathbb{E}_{\mathcal{D}}[\mathrm{tr}\left((\mathfrak{Z}_1-\bar{\mathfrak{Z}}_1)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)]\\ \vdots \\ \mathbb{E}_{\mathcal{D}}[\mathrm{tr}\left((\mathfrak{Z}_P-\bar{\mathfrak{Z}}_P)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)]\end{pmatrix}. \tag{F.48}$$

Using the same steps as in the proof of Lemma 5 in Shi and Lee (2017), it can be shown that $\frac{1}{\sqrt{nT}}||\boldsymbol{l}_{3.1}||_2^2=o_P(1)$ and then, finally, by using the explicit expressions for $\mathfrak{Z}_p-\bar{\mathfrak{Z}}_p$ given in the proof of Lemmas F.1(v),

$$\frac{1}{\sqrt{nT}}\begin{pmatrix}\mathbb{E}_{\mathcal{D}}[\mathrm{tr}\left((\mathfrak{Z}_1-\bar{\mathfrak{Z}}_1)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)]\\ \vdots \\ \mathbb{E}_{\mathcal{D}}[\mathrm{tr}\left((\mathfrak{Z}_P-\bar{\mathfrak{Z}}_P)'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)]\end{pmatrix}=\frac{1}{\sqrt{nT}}\begin{pmatrix}\sum_{t=1}^{T-1}\mathrm{tr}(\boldsymbol{J}_0\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{J}_h')\mathrm{tr}(\boldsymbol{W}_1\boldsymbol{A}^h\boldsymbol{S}^{-1})\\ \vdots \\ \sum_{t=1}^{T-1}\mathrm{tr}(\boldsymbol{J}_0\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{J}_h')\mathrm{tr}(\boldsymbol{W}_Q\boldsymbol{A}^h\boldsymbol{S}^{-1})\\ \boldsymbol{0}_{K^*\times 1}\\ \sum_{h=1}^{T-1}\mathrm{tr}(\boldsymbol{J}_0\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{J}_h')\mathrm{tr}(\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1})\\ \sum_{h=1}^{T-1}\mathrm{tr}(\boldsymbol{J}_0\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{J}_h')\mathrm{tr}(\boldsymbol{W}_1\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1})\\ \vdots \\ \sum_{h=1}^{T-1}\mathrm{tr}(\boldsymbol{J}_0\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{J}_h')\mathrm{tr}(\boldsymbol{W}_Q\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1})\end{pmatrix}, \tag{F.49}$$

which, combined with the previous parts, yields the result. □

Proof of Lemma A.6. In this proof a central limit theorem is proven for $\mathbb{S}\boldsymbol{c}$. The steps are similar to the proof of Lemma 13 in Yu et al. (2008), with modifications due to the increasing number of parameters. Let $\boldsymbol{v}\in\mathbb{R}^L$ with elements v_l and $||\boldsymbol{v}||_2$ bounded for all L. Also, recall the definition of $\mathbb{S}$ given in Assumption 7.1. Then $\boldsymbol{v}'\mathbb{S}\boldsymbol{c}$ equals

$$\boldsymbol{v}'\mathbb{S}\left(\mathbb{Z}'\mathrm{vec}(\boldsymbol{\varepsilon})+\begin{pmatrix}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{G}_1^*\boldsymbol{\varepsilon})\\ \vdots \\ \mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{G}_Q^*\boldsymbol{\varepsilon})\\ \boldsymbol{0}_{K+1}\end{pmatrix}\right)$$

$$=\boldsymbol{v}'\mathbb{S}\left(\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_1\boldsymbol{M}_{\boldsymbol{F}^0}+(\mathfrak{Z}_1-\bar{\mathfrak{Z}}_1))\boldsymbol{\varepsilon}'\right)\\ \vdots \\ \mathrm{tr}\left((\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_P\boldsymbol{M}_{\boldsymbol{F}^0}+(\mathfrak{Z}_P-\bar{\mathfrak{Z}}_P))\boldsymbol{\varepsilon}'\right)\end{pmatrix}+\begin{pmatrix}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{G}_1^*\boldsymbol{\varepsilon})\\ \vdots \\ \mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{G}_Q^*\boldsymbol{\varepsilon})\\ \boldsymbol{0}_{K+1}\end{pmatrix}\right)$$

$$
= \boldsymbol{v}'\mathbb{S}\left(\begin{pmatrix}\mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_1\boldsymbol{M}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')\\ \vdots \\ \mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_P\boldsymbol{M}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')\end{pmatrix} + \begin{pmatrix}\sum_{h=1}^{\infty}\mathrm{tr}(\boldsymbol{W}_1\boldsymbol{A}^h\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_h^*\boldsymbol{\varepsilon}')\\ \vdots \\ \sum_{h=1}^{\infty}\mathrm{tr}(\boldsymbol{W}_Q\boldsymbol{A}^h\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_h^*\boldsymbol{\varepsilon}')\\ \mathbf{0}_{K^*\times 1}\\ \sum_{h=1}^{\infty}\mathrm{tr}(\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_h^*\boldsymbol{\varepsilon}')\\ \sum_{h=1}^{\infty}\mathrm{tr}(\boldsymbol{W}_1\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_h^*\boldsymbol{\varepsilon}')\\ \vdots \\ \sum_{h=1}^{\infty}\mathrm{tr}(\boldsymbol{W}_Q\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_h^*\boldsymbol{\varepsilon}')\end{pmatrix} + \begin{pmatrix}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{G}_1^*\boldsymbol{\varepsilon}) - T\sigma_0^2\mathrm{tr}(\boldsymbol{G}_1^*)\\ \vdots \\ \mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{G}_Q^*\boldsymbol{\varepsilon}) - T\sigma_0^2\mathrm{tr}(\boldsymbol{G}_Q^*)\\ \mathbf{0}_{K\times 1}\end{pmatrix}\right), \tag{F.50}
$$

where the last line follows from applying the expressions for $\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p$ given in the proof of Lemma F.1(v), and also noticing $T\sigma_0^2\mathrm{tr}(\boldsymbol{G}_q^*) = T\sigma_0^2\mathrm{tr}(\boldsymbol{G}_q - \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_q)\boldsymbol{I}_n) = 0$. Now, define the matrices

$$
\mathfrak{D} := \sum_{l=1}^{L}\sum_{p=1}^{P} v_l(\mathbb{S})_{lp}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_p\boldsymbol{M}_{\boldsymbol{F}^0} \tag{F.51}
$$

$$
\mathfrak{U} := \sum_{l=1}^{L}\sum_{p=1}^{P}\sum_{h=1}^{\infty} v_l(\mathbb{S})_{lp}\boldsymbol{\mathcal{A}}_p\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_{h+1}^* \tag{F.52}
$$

$$
\mathfrak{B} := \sum_{l=1}^{L}\sum_{p=1}^{P} v_l(\mathbb{S})_{lp}\frac{1}{2}\left(\boldsymbol{\mathcal{G}}_p + \boldsymbol{\mathcal{G}}_p'\right) \tag{F.53}
$$

with elements d_{it}, u_{it} and b_{ij} respectively, where

$$
\boldsymbol{\mathcal{A}}_p := \begin{cases}\boldsymbol{W}_q\boldsymbol{A} & \text{for } p = 1, ..., Q \text{ with } q = p\\ \mathbf{0}_{n\times n} & \text{for } p = Q+1, ..., Q+K^*\\ \boldsymbol{I}_n & \text{for } p = Q+K^*+1\\ \boldsymbol{W}_q & \text{for } p = Q+K^*+2, ..., P \text{ with } q = p-Q-K^*-1,\end{cases}, \tag{F.54}
$$

and $\boldsymbol{\mathcal{G}}_p := \boldsymbol{G}_p^*$ for $p = 1, ..., Q$, and $\mathbf{0}_{n\times n}$ for $p = Q+1, ..., P$. Using the matrices defined in (F.51), (F.52) and (F.53), let

$$
\mathcal{J} := \boldsymbol{v}'\mathbb{S}\boldsymbol{c} = \sum_{t=1}^{T}\sum_{i=1}^{n}\left((u_{it-1} + d_{it})\varepsilon_{it} + b_{ii}(\varepsilon_{it}^2 - \sigma_0^2) + 2\sum_{j=1}^{i-1} b_{ij}\varepsilon_{jt}\varepsilon_{it}\right) =: \sum_{t=1}^{T}\sum_{i=1}^{n} j_{it}, \tag{F.55}
$$

and the variance of $\mathcal{J}$ be denoted $\sigma_{\mathcal{J}}^2$. In what follows the aim is to show that the standardised sum $\frac{\mathcal{J}}{\sigma_{\mathcal{J}}} \xrightarrow{d} \mathcal{N}(0,1)$ for any $\boldsymbol{v}$. Following Yu et al. (2008), define the σ-algebra

$\mathcal{F}_{it} := \sigma(\varepsilon_{11}, ..., \varepsilon_{n1}, \varepsilon_{12}, ..., \varepsilon_{n2}, \varepsilon_{1t}, ..., \varepsilon_{it})$. A central limit theorem for martingale difference arrays applies to $\frac{\mathcal{J}}{\sigma_{\mathcal{J}}}$ if the following two conditions are met:

$$\frac{1}{\sigma_{\mathcal{J}}^{2+\delta}} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbb{E}[|j_{it}|^{2+\delta}] \to 0, \tag{F.56}$$

$$\frac{1}{\sigma_{\mathcal{J}}^{2}} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbb{E}[j_{it}^2 | \mathcal{F}_{i-1t}] \xrightarrow{P} 1. \tag{F.57}$$

Conditions (F.56) and (F.57) are now demonstrated in turn. Beginning with (F.56), note that

$$\begin{aligned}|j_{it}| &\leq |u_{it-1} + d_{it}||\varepsilon_{it}| + |b_{ii}||\varepsilon_{it}^2 + \sigma_0^2| + 2\sum_{j=1}^{i-1} |b_{ij}||\varepsilon_{jt}||\varepsilon_{it}| \\ &= |u_{it-1} + d_{it}||\varepsilon_{it}| + |b_{ii}|^{\frac{1}{p}}|b_{ii}|^{\frac{1}{q}}|\varepsilon_{it}^2 + \sigma_0^2| + 2\sum_{j=1}^{i-1} |b_{ij}|^{\frac{1}{p}}|b_{ij}|^{\frac{1}{q}}|\varepsilon_{jt}||\varepsilon_{it}|\end{aligned}$$

for any $p, q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$. By Hölder's inequality

$$\begin{aligned}&|u_{it-1} + d_{it}||\varepsilon_{it}| + |b_{ii}|^{\frac{1}{p}}|b_{ii}|^{\frac{1}{q}}|\varepsilon_{it}^2 + \sigma_0^2| + 2\sum_{j=1}^{i-1} |b_{ij}|^{\frac{1}{p}}|b_{ij}|^{\frac{1}{q}}|\varepsilon_{jt}||\varepsilon_{it}| \\ &= |u_{it-1} + d_{it}||\varepsilon_{it}| + |b_{ii}|^{\frac{1}{p}}|b_{ii}|^{\frac{1}{q}}|\varepsilon_{it}^2 + \sigma_0^2| + 2|b_{i1}|^{\frac{1}{p}}|b_{i1}|^{\frac{1}{q}}|\varepsilon_{1t}||\varepsilon_{it}| \\ &\quad + \ ... \ + 2|b_{ii-1}|^{\frac{1}{p}}|b_{ii-1}|^{\frac{1}{q}}|\varepsilon_{i-1t}||\varepsilon_{it}| \\ &\leq \left(|u_{it-1} + d_{it}|^p + \sum_{j=1}^{i} |b_{ij}|\right)^{\frac{1}{p}} \left(|\varepsilon_{it}|^q + |b_{ii}||\varepsilon_{it}^2 - \sigma_0^2|^q + 2^q|\varepsilon_{it}|^q \sum_{j=1}^{i-1} |b_{ij}||\varepsilon_{jt}|^q\right)^{\frac{1}{q}}.\end{aligned}$$

Hence

$$\begin{aligned}\mathbb{E}\left[|j_{it}|^q\right] \leq \mathbb{E}&\left[\left(|u_{it-1} + d_{it}|^p + \sum_{j=1}^{i} |b_{ij}|\right)^{\frac{q}{p}}\right] \\ &\times \left(\mathbb{E}\left[|\varepsilon_{it}|^q\right] + |b_{ii}|\mathbb{E}\left[|\varepsilon_{it}^2 - \sigma_0^2|^q\right] + 2^q\mathbb{E}\left[|\varepsilon_{it}|^q\right] \sum_{j=1}^{i-1} |b_{ij}|\mathbb{E}\left[|\varepsilon_{jt}|^q\right]\right).\end{aligned} \tag{F.58}$$

Let $q = 2 + \delta$ with δ small. By Assumption 1.1, ε_{it} has finite fourth moments. With $q < 4$, then $\mathbb{E}[|\varepsilon_{it} - \sigma_0^2|^q] \leq c_1$, $\mathbb{E}\left[|\varepsilon_{it}|^q\right] \leq c_2$. Now, for the matrix $\boldsymbol{\mathfrak{B}}$.

$$||\boldsymbol{\mathfrak{B}}||_1 \leq \sum_{l=1}^{L} \sum_{p=1}^{P} |v_l||(\mathbb{S})_{lp}| \left|\left|\frac{1}{2}\left(\boldsymbol{\mathcal{G}}_p + \boldsymbol{\mathcal{G}}_p'\right)\right|\right|_1$$

$$
\begin{aligned}
&\leq \sum_{l=1}^{L} |v_l| \left(\sum_{p=1}^{P} |(\mathbb{S})_{lp}|^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^{P} ||\boldsymbol{\mathcal{G}}_p||_1^2 \right)^{\frac{1}{2}} \\
&\leq \left(\sum_{l=1}^{L} |v_l|^2 \right)^{\frac{1}{2}} \left(\sum_{l=1}^{L} \sum_{p=1}^{P} |(\mathbb{S})_{lp}|^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^{P} \left|\left| \frac{1}{2} \left(\boldsymbol{\mathcal{G}}_p + \boldsymbol{\mathcal{G}}_p' \right) \right|\right|_1^2 \right)^{\frac{1}{2}} \\
&\leq \sqrt{P} ||v||_2 ||\mathbb{S}||_F \max_{1 \leq p \leq P} \left|\left| \frac{1}{2} \left(\boldsymbol{\mathcal{G}}_p + \boldsymbol{\mathcal{G}}_p' \right) \right|\right|_1 = O(\sqrt{P}),
\end{aligned}
$$

since $||\boldsymbol{v}||_2 \leq c$, $||\mathbb{S}||_F \leq \sqrt{\text{rank}(\mathbb{S})} ||\mathbb{S}||_2$ with $\text{rank}(\mathbb{S}) \leq L$, $||\mathbb{S}||_2 < c$ by Assumption 7.1, and using the fact that $\boldsymbol{\mathcal{G}}_p$ is UB by Assumptions 2.2 and 2.3. Similarly, it can be shown that $||\boldsymbol{\mathcal{B}}||_\infty = O(\sqrt{P})$. Therefore returning to (F.58)

$$
\mathbb{E}\left[|j_{it}|^q\right] \leq \mathbb{E}\left[\left((|u_{it-1}| + |d_{it}|)^p + \sum_{j=1}^{i} |b_{ij}| \right)^{\frac{q}{p}} \right] O(\sqrt{P}).
$$

Next, by the c_r inequality (see, for instance Davidson, 1994, result 9.28), and since $\frac{q}{p} = 1+\delta$,

$$
\begin{aligned}
\mathbb{E}\left[\left((|u_{it-1}| + |d_{it}|)^p + \sum_{j=1}^{i} |b_{ij}| \right)^{\frac{q}{p}} \right] &\leq 2^{\frac{q}{p}-1} \left(\mathbb{E}[(|u_{it-1}| + |d_{it}|)^q] + \left(\sum_{j=1}^{i} |b_{ij}| \right)^{\frac{q}{p}} \right) \\
&\leq 2^{\frac{q}{p}-1} \left(\mathbb{E}[(|u_{it-1}| + |d_{it}|)^q] + O(P^{\frac{q}{2p}}) \right) \\
&\leq 2^{\frac{q}{p}-1} \left(2^{q-1} (\mathbb{E}[|u_{it-1}|^q] + \mathbb{E}[|d_{it}|^q) + O(P^{\frac{q}{2p}}) \right).
\end{aligned} \tag{F.59}
$$

Now,

$$
\begin{aligned}
|u_{it-1}| &\leq \sum_{l=1}^{L} \sum_{p=1}^{P} |v_l| |(\mathbb{S})_{lp}| \left| \left(\sum_{h=1}^{\infty} \boldsymbol{\mathcal{A}}_p \boldsymbol{A}^{h-1} \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_{h+1}^* \right)_{it-1} \right| \\
&\leq ||\boldsymbol{v}||_2 ||\mathbb{S}||_F \left(\sum_{p=1}^{P} \left| \left(\sum_{h=1}^{\infty} \boldsymbol{\mathcal{A}}_p \boldsymbol{A}^{h-1} \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_{h+1}^* \right)_{it-1} \right|^2 \right)^{\frac{1}{2}} \\
&\leq ||\boldsymbol{v}||_2 ||\mathbb{S}||_F \sqrt{P} \operatorname*{arg\,max}_{1 \leq p \leq P} \left| \left(\sum_{h=1}^{\infty} \boldsymbol{\mathcal{A}}_p \boldsymbol{A}^{h-1} \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_{h+1}^* \right)_{it} \right|.
\end{aligned} \tag{F.60}
$$

Hence

$$
\mathbb{E}[|u_{it-1}|^q] \leq ||\boldsymbol{v}||_2^q ||\mathbb{S}||_F^q P^{\frac{q}{2}} \mathbb{E}\left[\left(\operatorname*{arg\,max}_{1 \leq p \leq P} \left| \left(\sum_{h=1}^{\infty} \boldsymbol{\mathcal{A}}_p \boldsymbol{A}^{h-1} \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_{h+1}^* \right)_{it} \right| \right)^q \right]. \tag{F.61}
$$

Recall that $q = 2+\delta$. By the same steps as in Lemma 10 in Yu et al. (2008), it can be shown that $\mathbb{E}\big[\big(\big|\big(\sum_{h=1}^{\infty}\boldsymbol{\mathcal{A}}_p\boldsymbol{A}^{h-1}\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}^*_{h+1}\big)_{it}\big|\big)^4\big]$ is $O(1)$, uniformly across i,t, and p. Therefore $\mathbb{E}[|u_{it-1}|^q] = O(P^{\frac{q}{2}}) = O(P^{1+\frac{\delta}{2}})$. By similar steps,

$$
\begin{aligned}
|d_{it}| &\leq \sum_{l=1}^{L}\sum_{p=1}^{P}|v_l||(\mathbb{S})_{lp}||\left(\boldsymbol{M}_{\boldsymbol{F}^0}\bar{\boldsymbol{\mathcal{Z}}}_p\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right)_{it}| \\
&\leq ||\boldsymbol{v}||_2||\mathbb{S}||_F\left(\sum_{p=1}^{P}\left(\left(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\boldsymbol{\mathcal{Z}}}_p\boldsymbol{M}_{\boldsymbol{F}^0}\right)_{it}\right)^2\right)^{\frac{1}{2}}. \qquad \text{(F.62)}
\end{aligned}
$$

Hence

$$
\mathbb{E}\left[|d_{it}|^q\right] \leq ||\boldsymbol{v}||_2^q||\mathbb{S}||_F^2\mathbb{E}\left[\left(\sum_{p=1}^{P}\left(\left(\boldsymbol{M}_{\boldsymbol{F}^0}\bar{\boldsymbol{\mathcal{Z}}}_p\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right)_{it}\right)^2\right)^{\frac{q}{2}}\right]. \qquad \text{(F.63)}
$$

It is straightforward to see that $\mathbb{E}\big[\big(\sum_{p=1}^{P}\left(\left(\boldsymbol{M}_{\boldsymbol{F}^0}\bar{\boldsymbol{\mathcal{Z}}}_p\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right)_{it}\right)^2\big)^2\big]$ is $O(P^2)$ and hence $\mathbb{E}\left[|d_{it}|^q\right] = O(P^2)$. Returning to (F.58),

$$
\begin{aligned}
\mathbb{E}\left[|j_{it}|^q\right] &\leq \left(O(P^{1+\frac{\delta}{2}}) + O(P^2) + O(P^{\frac{q}{2p}})\right)O(\sqrt{P}) \\
&= O(P^{\frac{3+\delta}{2}}), \qquad \text{(F.64)}
\end{aligned}
$$

therefore

$$
\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[|j_{it}^{2+\delta}|] = O(nTP^{\frac{5}{2}}), \qquad \text{(F.65)}
$$

and

$$
\frac{1}{\sigma_{\mathcal{J}}^{2+\delta}}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[|j_{it}^{2+\delta}|] = \frac{1}{\left(\frac{\sigma_{\mathcal{J}}^{2+\delta}}{(nT)^{1+\frac{\delta}{2}}}\right)}\frac{1}{(nT)^{1+\frac{\delta}{2}}}O(nTP^{\frac{5}{2}}) = \frac{1}{\left(\frac{\sigma_{\mathcal{J}}^{2+\delta}}{(nT)^{1+\frac{\delta}{2}}}\right)}\frac{1}{(nT)^{\frac{\delta}{2}}}O(P^{\frac{5}{2}}). \qquad \text{(F.66)}
$$

Since $\left(\frac{\sigma_{\mathcal{J}}^{2+\delta}}{(nT)^{1+\frac{\delta}{2}}}\right)^{-1} = O(1)$,

$$
\frac{1}{\sigma_{\mathcal{J}}^{2+\delta}}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[|j_{it}^{2+\delta}|] = O\left(\frac{P^{\frac{5}{2}}}{(nT)^{\frac{\delta}{2}}}\right) = o(1) \qquad \text{(F.67)}
$$

for δ sufficiently large under Assumption 6.1. This verifies (F.56). Next Condition (F.57) is established. Note that

$$\frac{1}{\sigma_{\mathcal{J}}^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[j_{it}^2|\mathcal{F}_{i-1t}] \xrightarrow{P} 1$$

is equivalent to

$$\frac{1}{\sigma_{\mathcal{J}}^2}\left(\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[j_{it}^2|\mathcal{F}_{i-1t}] - \sigma_{\mathcal{J}}^2\right) \xrightarrow{P} 0$$

and hence also to

$$\frac{1}{\sigma_{\mathcal{J}}^2}\left(\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[j_{it}^2|\mathcal{F}_{i-1t}] - \sigma_{\mathcal{J}}^2\right) = o_P(1). \tag{F.68}$$

First,

$$\begin{aligned}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[j_{it}^2|\mathcal{F}_{i-1t}] = \sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}\Bigg[&(u_{it-1}+d_{it})^2\varepsilon_{it}^2 + (u_{it-1}+d_{it})\varepsilon_{it}b_{ii}(\varepsilon_{it}^2-\sigma_0^2)\\ &+ 2(u_{it-1}+d_{it})\varepsilon_{it}^2\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right) + b_{ii}(\varepsilon_{it}^2-\sigma_0^2)(u_{it-1}+d_{it})\varepsilon_{it} + b_{ii}^2(\varepsilon_{it}^2-\sigma_0^2)^2\\ &+ 2b_{ii}(\varepsilon_{it}^2-\sigma_0^2)\varepsilon_{it}\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right) + 2\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\varepsilon_{it}^2(u_{it-1}+d_{it})\\ &+ 2\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\varepsilon_{it}b_{ii}(\varepsilon_{it}^2-\sigma_0^2) + 4\varepsilon_{it}^2\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)|\mathcal{F}_{i-1t}\Bigg].\end{aligned}$$

Recall $\mathcal{M}_1^\varepsilon = 0$ and that $\mathcal{M}_2^\varepsilon = \sigma_0^2$. Moreover, ε_{jt} is independent of u_{it-1} and d_{it} for all $i, j = 1, ..., n$. Thus,

$$\begin{aligned}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[j_{it}^2|\mathcal{F}_{i-1t}] = \sum_{t=1}^{T}\sum_{i=1}^{n}&\sigma_0^2\mathbb{E}[(u_{it-1}+d_{it})^2] + 2\mathbb{E}[u_{it}+d_{it}]\mathcal{M}_3^\varepsilon b_{ii}\\ &+ 4\mathbb{E}[u_{it-1}+d_{it}]\sigma_0^2\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right) + b_{ii}^2\mathcal{M}_4^\varepsilon - b_{ii}^2\sigma_0^4\\ &+ 4b_{ii}\mathcal{M}_3^\varepsilon\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right) + 4\sigma_0^2\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right). \end{aligned}\tag{F.69}$$

Next, since $\mathbb{E}[\mathcal{J}] = 0$,

$$\sigma_{\mathcal{J}}^2 = \sum_{t=1}^{T}\sum_{i=1}^{n}\sigma_0^2\mathbb{E}[(u_{it-1}+d_{it})^2] + 2\mathbb{E}[u_{it}+d_{it}]\mathcal{M}_3^\varepsilon b_{ii} + b_{ii}^2\mathcal{M}_4^\varepsilon - b_{ii}^2\sigma_0^4 + 4\sigma_0^4\left(\sum_{j=1}^{i-1}b_{ij}^2\right). \tag{F.70}$$

Combining (F.69) and (F.70),

$$\begin{aligned}
\frac{1}{\sigma_{\mathcal{J}}^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[j_{it}^2|\mathcal{F}_{i-1t}]-\sigma_{\mathcal{J}}^2 &= \frac{1}{\sigma_{\mathcal{J}}^2}\sum_{t=1}^{T}\sum_{i=1}^{n}4\left(\mathbb{E}[u_{it-1}+d_{it}]\sigma_0^2+b_{ii}\mathcal{M}_3^{\varepsilon}\right)\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\\
&\quad+4\sigma_0^2\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)-4\sigma_0^4\left(\sum_{j=1}^{i-1}b_{ij}^2\right)\\
&= \frac{1}{\sigma_{\mathcal{J}}^2}\sum_{t=1}^{T}\sum_{i=1}^{n}4\left(\mathbb{E}[u_{it-1}+d_{it}]\sigma_0^2+b_{ii}\mathcal{M}_3^{\varepsilon}\right)\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\\
&\quad+8\sigma_0^2\sum_{j=1}^{i-1}\sum_{l=1}^{j-1}b_{ij}b_{il}\varepsilon_{jt}\varepsilon_{lt}+4\sigma_0^2\sum_{j=1}^{i-1}b_{ij}^2(\varepsilon_{jt}^2-\sigma_0^2)\\
&= \frac{4\sigma_0^2}{\frac{1}{nT}\sigma_{\mathcal{J}}^2}\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\mathbb{E}[u_{it-1}+d_{it}]+b_{ii}\frac{\mathcal{M}_3^{\varepsilon}}{\sigma_0^2}\right)\left(\sum_{j=1}^{i-1}b_{ij}\varepsilon_{jt}\right)\\
&\quad+2\sum_{j=1}^{i-1}\sum_{l=1}^{j-1}b_{ij}b_{il}\varepsilon_{jt}\varepsilon_{lt}+\sum_{j=1}^{i-1}b_{ij}^2(\varepsilon_{jt}^2-\sigma_0^2)\\
&=: \frac{4\sigma_0^2}{\frac{1}{nT}\sigma_{\mathcal{J}}^2}(H_1+2H_2+H_3). \qquad \text{(F.71)}
\end{aligned}$$

With $\frac{1}{nT}\sigma_{\mathcal{J}}^2>0$, it remains only to be shown that H_1, H_2 and H_3 are $o_P(1)$. It is straightforward to see that $\mathbb{E}[H_1]=\mathbb{E}[H_2]=\mathbb{E}[H_3]=0$. Next,

$$\begin{aligned}
\mathbb{E}[H_1^2] &= \mathbb{E}\left[\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{j=1}^{n-1}\left(\sum_{i=j+1}^{n}\left(\mathbb{E}[u_{it-1}+d_{it}]+b_{ii}\frac{\mathcal{M}_3^{\varepsilon}}{\sigma_0^2}\right)b_{ij}\varepsilon_{jt}\right)\right)^2\right]\\
&= \frac{\sigma_0^2}{(nT)^2}\sum_{t=1}^{T}\sum_{j=1}^{n-1}\left(\sum_{i=j+1}^{n}\left(\mathbb{E}[u_{it-1}+d_{it}]+b_{ii}\frac{\mathcal{M}_3^{\varepsilon}}{\sigma_0^2}\right)b_{ij}\right)^2\\
&\le \frac{\sigma_0^2}{(nT)^2}\sum_{t=1}^{T}\sum_{j=1}^{n-1}\left(\max_{1\le i\le n}\left\{\left|\mathbb{E}[u_{it-1}+d_{it}]+b_{ii}\frac{\mathcal{M}_3^{\varepsilon}}{\sigma_0^2}\right|\right\}\sum_{i=j+1}^{n}|b_{ij}|\right)^2\\
&= \frac{\sigma_0^2}{(nT)^2}T\left(\max_{1\le i\le n}\left\{\left|\mathbb{E}[u_{it-1}+d_{it}]+b_{ii}\frac{\mathcal{M}_3^{\varepsilon}}{\sigma_0^2}\right|\right\}\right)^2\sum_{j=1}^{n-1}\left(\sum_{i=j+1}^{n}|b_{ij}|\right)^2. \qquad \text{(F.72)}
\end{aligned}$$

Since $\mathbb{E}[|u_{it-1}|], \mathbb{E}[|d_{it}|], b_{ii} = O(\sqrt{P})$ and $\sum_{j=1}^{n-1}\left(\sum_{i=j+1}^{n}|b_{ij}|\right)^2 \leq \sum_{j=1}^{n}\left(\sum_{i=1}^{n}|b_{ij}|\right)^2 = O(nP)$, then $\mathbb{E}[H_1^2] = O(\frac{P}{nT}) = o(1)$. For $\mathbb{E}[2H_2^2]$,

$$
\begin{aligned}
\mathbb{E}[2H_2^2] &= \mathbb{E}\left[\left(\frac{2}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{l=1}^{j-1} b_{ij}b_{il}\varepsilon_{jt}\varepsilon_{lt}\right)^2\right] \\
&= \frac{4}{(nT)^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{l=1}^{j-1}\sum_{t'=1}^{T}\sum_{i'=1}^{n}\sum_{j'=1}^{i'-1}\sum_{l'=1}^{j'-1} b_{ij}b_{il}b_{i'j'}b_{i'l'}\mathbb{E}\left[\varepsilon_{jt}\varepsilon_{lt}\varepsilon_{j't'}\varepsilon_{l't'}\right] \\
&= \frac{4\sigma_0^4}{n^2T}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{l=1}^{j-1}\sum_{m=j+1}^{n} b_{ij}b_{il}b_{mj}b_{ml} \\
&\leq \frac{4\sigma_0^4}{n^2T}\sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|\sum_{m=1}^{n}|b_{mj}|\sum_{l=1}^{n}|b_{il}b_{ml}| \\
&\leq \frac{4\sigma_0^4}{n^2T}\sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|\sum_{m=1}^{n}|b_{mj}|\max_{1\leq l\leq n}\{|b_{il}|\}\sum_{l=1}^{n}|b_{ml}| \\
&= O\left(\frac{P^2}{nT}\right) = o(1)
\end{aligned} \tag{F.73}
$$

as $\sum_{j=1}^{n}|b_{ij}| = O(\sqrt{P})$ for $i = 1, ..., n$. Finally,

$$
\begin{aligned}
\mathbb{E}[H_3^2] &= \mathbb{E}\left[\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{i-1} b_{ij}^2(\varepsilon_{jt}^2 - \sigma_0^2)\right)^2\right] \\
&= \frac{1}{(nT)^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{t'=1}^{T}\sum_{i'=1}^{n}\sum_{j'=1}^{i-1} b_{ij}^2 b_{i'j'}^2\mathbb{E}[\varepsilon_{jt}^2\varepsilon_{j't'}^2] - b_{ij}^2 b_{i'j'}^2\sigma_0^4.
\end{aligned} \tag{F.74}
$$

For $j \neq j'$ or $t \neq t'$, $\mathbb{E}[\varepsilon_{jt}^2\varepsilon_{j't'}^2] = \sigma_0^2$ and for $j = j'$ and $t = t'$ $\mathbb{E}[\varepsilon_{jt}^2\varepsilon_{j't'}^2] = \mathcal{M}_4^{\varepsilon}$ and hence

$$
\begin{aligned}
\mathbb{E}[H_3^2] &= \frac{1}{n^2T}(\mathcal{M}_4^{\varepsilon} - \sigma_0^4)\sum_{j=1}^{n-1}\left(\sum_{i=j+1}^{i-1} b_{ij}\right)^2 \\
&\leq \frac{1}{n^2T}(\mathcal{M}_4^{\varepsilon} - \sigma_0^4)\max_{1\leq i,j\leq n}|b_{ij}|^2\sum_{j=1}^{n-1}\left(\sum_{i=j+1}^{i-1}|b_{ij}|\right)^2 \\
&= O\left(\frac{P^2}{nT}\right) = o(1).
\end{aligned} \tag{F.75}
$$

Given $\mathbb{E}[H_1^2], \mathbb{E}[H_2^2]$ and $\mathbb{E}[H_3^2]$ are $o\,(1)$, then, by Chebyshev's inequality, H_1, H_2 and H_3 are $o_P(1)$. This verifies Condition (F.57) and it follows that $\frac{\mathcal{J}}{\sigma_{\mathcal{J}}} \xrightarrow{d} \mathcal{N}(0,1)$. In the remainder of this proof it is shown that this is equivalent to the expression given in the statement of the lemma. Recall (F.70),

$$
\begin{aligned}
\sigma^2_{\mathcal{J}} &= \sum_{t=1}^{T}\sum_{i=1}^{n} \sigma_0^2\mathbb{E}[(u_{it-1}+d_{it})^2] + 2\mathbb{E}[u_{it}+d_{it}]\mathcal{M}_3^{\varepsilon}b_{ii} + b_{ii}^2(\mathcal{M}_4^{\varepsilon}-\sigma_0^4) + 4\sigma_0^4\left(\sum_{j=1}^{i-1} b_{ij}^2\right) \\
&= \left(\sum_{t=1}^{T}\sum_{i=1}^{n} 2\sigma_0^4 b_{ii}^2 + 4\sigma_0^2\left(\sum_{j=1}^{i-1} b_{ij}^2\right) + \sigma_0^2\mathbb{E}[(u_{it-1}+d_{it})^2]\right) \\
&\quad + \left(\sum_{t=1}^{T}\sum_{i=1}^{n} 2\mathcal{M}_{\varepsilon}^3\mathbb{E}[u_{it-1}+d_{it}]b_{ii} + (\mathcal{M}_{\varepsilon}^4 - 3\sigma_0^4)b_{ii}^2\right) \\
&=: l_1 + l_2.
\end{aligned}
\tag{F.76}
$$

Rearrange l_1 to give

$$
\begin{aligned}
l_1 &= 2\sigma_0^2\sum_{t=1}^{T}\sum_{i=1}^{n}\left(b_{ii}^2 + 2\left(\sum_{j=1}^{i-1} b_{ij}^2\right)\right) + \sigma_0^2\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[(u_{it-1}+d_{it})^2] \\
&= 2T\sigma_0^4\mathrm{tr}(\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}) + \sigma_0^2\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[(u_{it-1}+d_{it})^2].
\end{aligned}
\tag{F.77}
$$

For the first term of (F.77),

$$
\begin{aligned}
2T\sigma_0^4\mathrm{tr}(\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}) &= 2T\sigma_0^4\mathrm{tr}\left(\sum_{l=1}^{L}\sum_{p=1}^{P}\sum_{l'=1}^{L}\sum_{p'=1}^{P} v_l(\mathbb{S})_{lp}\frac{1}{2}\left(\boldsymbol{\mathcal{G}}_p+\boldsymbol{\mathcal{G}}_p'\right)\frac{1}{2}\left(\boldsymbol{\mathcal{G}}_{p'}+\boldsymbol{\mathcal{G}}_{p'}'\right)(\mathbb{S})_{l'p'}v_{l'})\right) \\
&= nT\sigma_0^4\boldsymbol{v}'\mathbb{S}\boldsymbol{\Omega}\mathbb{S}'\boldsymbol{v},
\end{aligned}
\tag{F.78}
$$

using the definition of $\boldsymbol{\mathcal{G}}_p$ and $\boldsymbol{G}_q^*$. For the second term of (F.77), recalling that $\mathbb{Z}_p \coloneqq \boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\boldsymbol{\mathcal{Z}}}_p\boldsymbol{M}_{\boldsymbol{F}^0} + (\boldsymbol{\mathcal{Z}}_p - \bar{\boldsymbol{\mathcal{Z}}}_p)$,

$$
\begin{aligned}
&\sigma_0^2\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}[(u_{it-1}+d_{it})^2] \\
&= \sigma_0^2\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{l=1}^{L}\sum_{p=1}^{P} v_l(\mathbb{S})_{lp}(\mathbb{Z})_{pit}\right)^2\right] \\
&= \sigma_0^2\boldsymbol{v}'\mathbb{S}\mathbb{E}\left[\begin{pmatrix} \mathrm{tr}\left(\mathbb{Z}_1'\mathbb{Z}_1\right) & \cdots & \mathrm{tr}\left(\mathbb{Z}_1'\mathbb{Z}_P\right) \\ \vdots & \ddots & \vdots \\ \mathrm{tr}\left(\mathbb{Z}_P'\mathbb{Z}_1\right) & \cdots & \mathrm{tr}\left(\mathbb{Z}_P'\mathbb{Z}_P\right)\end{pmatrix}\right]\mathbb{S}'\boldsymbol{v}
\end{aligned}
$$

$$= \sigma_0^4 nT\boldsymbol{v}'\mathbb{S}\mathbb{E}\left[\left(\frac{1}{\sigma_0^2}\frac{1}{nT}\begin{pmatrix} \mathrm{tr}\,(\mathbb{Z}_1'\mathbb{Z}_1) & \cdots & \mathrm{tr}\,(\mathbb{Z}_1'\mathbb{Z}_P) \\ \vdots & \ddots & \vdots \\ \mathrm{tr}\,(\mathbb{Z}_P'\mathbb{Z}_1) & \cdots & \mathrm{tr}\,(\mathbb{Z}_P'\mathbb{Z}_P) \end{pmatrix} - \frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}\right)\right]\mathbb{S}'\boldsymbol{v}$$

$$+ \sigma_0^4 nT\boldsymbol{v}'\mathbb{S}\mathbb{E}\left[\frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}\right]\mathbb{S}'\boldsymbol{v}$$

$$= \sigma_0^4 nT\boldsymbol{v}'\mathbb{S}\mathbb{E}\left[\frac{1}{\sigma_0^2}\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}\right]\mathbb{S}'\boldsymbol{v} + o(1) \tag{F.79}$$

using Lemma F.2(ix). Therefore

$$l_1 = \sigma_0^4 nT\boldsymbol{v}'\mathbb{S}\mathbb{D}\mathbb{S}'\boldsymbol{v} + o(1). \tag{F.80}$$

Next consider l_2:

$$\begin{aligned} l_2 &= \sum_{t=1}^{T}\sum_{i=1}^{n} 2\mathcal{M}_\varepsilon^3\mathbb{E}[u_{it-1}+d_{it}]b_{ii} + \sum_{t=1}^{T}\sum_{i=1}^{n}(\mathcal{M}_\varepsilon^4 - 3\sigma_0^4)b_{ii}^2 \\ &=: l_{2.1} + l_{2.2} \end{aligned} \tag{F.81}$$

First, $l_{2.1}$ equals

$$2\mathcal{M}_\varepsilon^3\boldsymbol{v}'\mathbb{S}\underbrace{\begin{pmatrix}\begin{pmatrix} \sum_{t=1}^{T}\sum_{i=1}^{n}(\boldsymbol{G}_1^*)_{ii}\mathbb{E}\left[(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_1\boldsymbol{M}_{\boldsymbol{F}^0})_{it}\right] & \cdots & \sum_{t=1}^{T}\sum_{i=1}^{n}(\boldsymbol{G}_1^*)_{ii}\mathbb{E}\left[(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_P\boldsymbol{M}_{\boldsymbol{F}^0})_{it}\right] \\ \vdots & \ddots & \vdots \\ \sum_{t=1}^{T}\sum_{i=1}^{n}(\boldsymbol{G}_Q^*)_{ii}\mathbb{E}\left[(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_1\boldsymbol{M}_{\boldsymbol{F}^0})_{it}\right] & \cdots & \sum_{t=1}^{T}\sum_{i=1}^{n}(\boldsymbol{G}_Q^*)_{ii}\mathbb{E}\left[(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_P\boldsymbol{M}_{\boldsymbol{F}^0})_{it}\right] \end{pmatrix} \\ \mathbf{0}_{K\times P}\end{pmatrix}}_{\boldsymbol{\Phi}}\mathbb{S}'\boldsymbol{v}$$

$$= \mathcal{M}_\varepsilon^3\boldsymbol{v}'\mathbb{S}\mathbb{E}\left[\boldsymbol{\Phi}+\boldsymbol{\Phi}'\right]\mathbb{S}'\boldsymbol{v}, \tag{F.82}$$

since $\mathbb{E}\left[u_{it-1}\right] = 0$. Second,

$$l_{2.2} = (\mathcal{M}_\varepsilon^2 - 3\sigma_0^4)\boldsymbol{v}'\mathbb{S}\begin{pmatrix} \boldsymbol{\Xi} & \mathbf{0}_{Q\times K} \\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix}\mathbb{S}'\boldsymbol{v}, \tag{F.83}$$

and thus combining (F.82) and (F.83),

$$l_2 = nT\sigma_0^4\boldsymbol{v}'\mathbb{S}\mathbb{V}\mathbb{S}'\boldsymbol{v}. \tag{F.84}$$

Bringing together these results,

$$\sigma_{\mathcal{J}}^2 = nT\sigma_0^4\boldsymbol{v}'\mathbb{S}(\mathbb{D}+\mathbb{V})\mathbb{S}'\boldsymbol{v}, \tag{F.85}$$

and therefore,

$$\frac{\mathcal{J}}{\sigma_{\mathcal{J}}} = \frac{1}{\sqrt{nT}}\frac{1}{\sigma_0^2}\frac{\boldsymbol{v}'\mathbb{S}\boldsymbol{c}}{\sqrt{\boldsymbol{v}'\mathbb{S}(\mathbb{D}+\mathbb{V})\mathbb{S}'\boldsymbol{v}}}. \tag{F.86}$$

Since it has been established that $\frac{\mathcal{J}}{\sigma_{\mathcal{J}}} \xrightarrow{d} \mathcal{N}(0,1)$, then, by the Cramer-Wold device,

$$\frac{1}{\sqrt{nT}}\frac{1}{\sigma_0^2}\left(\mathbb{S}(\mathbb{D}+\mathbb{V})\mathbb{S}'\right)^{-\frac{1}{2}}\mathbb{S}\boldsymbol{c} \xrightarrow{d} \mathcal{N}(\boldsymbol{0}_{L\times L}, \boldsymbol{I}_L). \tag{F.87}$$

This completes the proof. □

G Proofs of Lemmas F.1–F.4

As in Appendix F, since the following intermediary lemmas are only used in proving Theorem 1 and Proposition 3, for notational simplicity, it is again assumed that all of the covariates are relevant, i.e. $K = K^0, Q = Q^0$, $P = P^0$ and $\boldsymbol{\mathcal{Z}}_{(1)} = \boldsymbol{\mathcal{Z}}$.

Proof of Lemma F.1(i). Recall that $\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}} = \boldsymbol{I}_R$. Then $\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}} = n\boldsymbol{I}_R$ and $||\hat{\boldsymbol{\Lambda}}||_2^2 = n$. Thus $||\hat{\boldsymbol{\Lambda}}||_2 = \sqrt{n}$. Next, $||\hat{\boldsymbol{\Lambda}}||_F^2 = \text{tr}(n\boldsymbol{I}_R) = Rn$, hence $||\hat{\boldsymbol{\Lambda}}||_F = \sqrt{Rn}$. □

Proof of Lemma F.1(ii).

$$\begin{aligned}
\mathbb{E}\left[||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'||_F^2\right] &= \sum_{r=1}^{R^0}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{t'=1}^{T}\mathbb{E}\left[f_{tr}^0 f_{t'r}^0 \varepsilon_{it}\varepsilon_{it'}\right] \\
&= \sum_{r=1}^{R^0}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{t'=1}^{T}\mathbb{E}\left[f_{tr}^0 f_{t'r}^0 \mathbb{E}_{\mathcal{D}}\left[\varepsilon_{it}\varepsilon_{it'}\right]\right] \\
&= \sigma_0^2\sum_{r=1}^{R^0}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbb{E}\left[(f_{tr}^0)^2\right] \\
&= O(nT),
\end{aligned}$$

using Assumptions 1 and 2.7, from which the first result follows. The remaining results are established similarly. □

Proof of Lemma F.1(iii). Using Lemma F.1(i) and the assumption that the matrices $\boldsymbol{G}_q$ are UB across q,

$$\left|\left|\left(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),\ldots,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\right)\right|\right|_2 \leq \left|\left|\left(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),\ldots,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\right)\right|\right|_F = \sqrt{\sum_{q=1}^{Q}||\text{vec}(\boldsymbol{G}_q'\hat{\boldsymbol{\Lambda}})||_2^2}$$

$$= \sqrt{\sum_{q=1}^{Q} ||\boldsymbol{G}_q'\hat{\boldsymbol{\Lambda}}||_F^2} \leq \sqrt{\sum_{q=1}^{Q} ||\boldsymbol{G}_q||_2^2 ||\hat{\boldsymbol{\Lambda}}||_F^2} \leq \sqrt{R}||\hat{\boldsymbol{\Lambda}}||_2 \sqrt{Q} \sqrt{\max_{1\leq q\leq Q} ||\boldsymbol{G}_q||_2^2} = O(\sqrt{Qn}),$$

under the normalisation $\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}} = \boldsymbol{I}_R$. □

Proof of Lemma F.1(iv). Follows by the same steps as in the proof of Lemma F.1(iii) and using Lemma A.2(vi). □

Proof of Lemma F.1(v). First, let $\boldsymbol{\varepsilon}_h^* := (\boldsymbol{\varepsilon}_{1-h}, ..., \boldsymbol{\varepsilon}_{T-h})$, which is the $n \times T$ matrix of lagged error terms. By recursive substitution of the model, explicit expressions for $\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p$ can be obtained as

$$\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p = \begin{cases} \boldsymbol{W}_q \sum_{h=1}^{\infty} \boldsymbol{A}^h \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^* & \text{for } p = 1, ..., Q \text{ with } q = p \\ \boldsymbol{0}_{n\times T} & \text{for } p = Q+1, ..., Q+K^* \\ \sum_{h=1}^{\infty} \boldsymbol{A}^{h-1} \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^* & \text{for } p = Q+K^*+1 \\ \boldsymbol{W}_q \sum_{h=1}^{\infty} \boldsymbol{A}^{h-1} \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^* & \text{for } p = Q+K^*+2, ..., P \text{ with } q = p-Q-K^*-1 \end{cases} \tag{G.1}$$

Using these expressions, for $p = 1, ..., Q$,

$$||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2 = \left|\left| \boldsymbol{W}_q \sum_{h=1}^{\infty} \boldsymbol{A}^h \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^* \right|\right|_2^2 \leq ||\boldsymbol{W}_q||_2^2 \left|\left| \sum_{h=1}^{T} \boldsymbol{A}^h \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^* + \boldsymbol{r} \right|\right|_2^2, \tag{G.2}$$

where $\boldsymbol{r} := \sum_{h=T+1}^{\infty} \boldsymbol{A}^h \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^*$. Let $\boldsymbol{\varepsilon}^{**} := (\boldsymbol{\varepsilon}_{1-T}, ..., \boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1, ..., \boldsymbol{\varepsilon}_T)$, an $n \times 2T$ matrix. Latala (2005) shows that under Assumption 1.1 $\mathbb{E}\left[||\boldsymbol{\varepsilon}^{**}||_2^2\right] = O(\max\{n, 2T\})$. Therefore, for the first term inside of the second norm on the right-hand side of (G.2),

$$\begin{aligned} \mathbb{E}\left[\left|\left| \sum_{h=1}^{T} \boldsymbol{A}^h \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^* \right|\right|_2^2\right] &\leq ||\boldsymbol{S}^{-1}||_2^2 \sum_{h=1}^{T} ||\boldsymbol{A}^h||_2^2 \mathbb{E}\left[||\boldsymbol{\varepsilon}_h^*||_2^2\right] \\ &\leq ||\boldsymbol{S}^{-1}||_2^2 \mathbb{E}\left[||\boldsymbol{\varepsilon}^{**}||_2^2\right] \sum_{h=1}^{T} (||\boldsymbol{A}||_2^2)^h = O(\max\{n, T\}), \end{aligned} \tag{G.3}$$

which follows from Assumption 2.3 and noting that $||\boldsymbol{\varepsilon}_h^*||_2 \leq ||\boldsymbol{\varepsilon}^{**}||_2$ since $\boldsymbol{\varepsilon}_h^*$ is a sub matrix of $\boldsymbol{\varepsilon}^{**}$. Also, $\mathbb{E}[||\boldsymbol{r}||_2^2] \leq \mathbb{E}[||\boldsymbol{r}||_F^2] = O(n)$ follows from the proof of Lemma 4(2) in Shi and Lee (2017). Now,

$$\mathbb{E}\left[||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2\right] \leq ||\boldsymbol{W}_q||_2^2 \mathbb{E}\left[\left(\left|\left| \sum_{h=1}^{T} \boldsymbol{A}^h \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}_h^* \right|\right|_2 + ||\boldsymbol{r}||_2\right)^2\right]. \tag{G.4}$$

Expanding the square in (G.4) it is straightforward to see that $\mathbb{E}\left[||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2\right] = O(\max\{n, T\})$. For $p = Q+1, ..., Q+K^*$, $\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p = 0$. For $p = Q+K^*+1, ..., P$, $\mathbb{E}[||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2] = O(\max\{n, T\})$, by similar arguments to those above. $\square$

Proof of Lemma F.1(vi).

$$\begin{aligned}
\mathbb{E}\left[\sum_{p=1}^{P} ||\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_F^2\right] &= \mathbb{E}\left[\sum_{p=1}^{P}\sum_{i=1}^{n}\sum_{i'=1}^{n}\sum_{t=1}^{T}\sum_{t'=1}^{T} \varepsilon_{it}\varepsilon_{it'}\bar{z}_{pi't}\bar{z}_{pi't'}\right] \\
&= \sigma_0^2 n \sum_{p=1}^{P}\sum_{i'=1}^{n}\sum_{t=1}^{T} \mathbb{E}\left[\bar{z}_{pi't}^2\right] = O(Pn^2T), \qquad \text{(G.5)}
\end{aligned}$$

since the elements of $\bar{\mathfrak{Z}}_p$ are independent of the error term and have finite second moments. $\square$

Proof of Lemma F.1(vii).

$$\begin{aligned}
\mathbb{E}\left[\sum_{p=1}^{P} ||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\bar{\mathfrak{Z}}_p||_F^2\right] &= \mathbb{E}\left[\sum_{p=1}^{P}\sum_{r=1}^{R^0}\sum_{t=1}^{T}\left(\sum_{\tau=1}^{T}\sum_{i=1}^{n} f_{\tau r}^0 \varepsilon_{i\tau}\bar{z}_{pit}\right)^2\right] \\
&= \mathbb{E}\left[\sum_{p=1}^{P}\sum_{r=1}^{R^0}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\sum_{\tau'=1}^{T}\sum_{j=1}^{n} f_{\tau r}^0 f_{\tau' r}^0 \bar{z}_{pit}\bar{z}_{pjt}\mathbb{E}_{\mathcal{D}}\left[\varepsilon_{i\tau}\varepsilon_{j\tau'}\right]\right] \\
&= \sigma_0^2 \sum_{p=1}^{P}\sum_{r=1}^{R^0}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} \mathbb{E}\left[(f_{\tau r}^0)^2\bar{z}_{pit}^2\right] \\
&\leq \sigma_0^2 \sum_{p=1}^{P}\sum_{r=1}^{R^0}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} \left(\mathbb{E}\left[(f_{\tau r}^0)^4\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\bar{z}_{pit}^4\right]\right)^{\frac{1}{2}} \\
&= O(PnT^2). \qquad \text{(G.6)}
\end{aligned}$$

This follows from Assumptions 1, 2.6 and 2.7 under which it can be shown that $\mathbb{E}[\bar{z}_{pjt}^4]$ is bounded uniformly in p, j and t. The second and third parts follow similarly. $\square$

Proof of Lemma F.1(viii).

$$\begin{aligned}
&\mathbb{E}\left[\left|\left|\frac{1}{T}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right|\right|_F^2\right] \\
&= \mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0\prime}\boldsymbol{\varepsilon}_t - \sigma_0^2\boldsymbol{\lambda}_r^{0\prime}\boldsymbol{\lambda}_s^0\right)^2\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{T^2}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_\tau\right] \\
&\quad - \frac{2}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_t\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\right] + \sigma_0^4\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\right] \\
&= \frac{1}{T^2}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_\tau\right] - \sigma_0^4\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\right]. \qquad \text{(G.7)}
\end{aligned}
$$

Using Lemma 3 in Yu et al. (2008),

$$
\begin{aligned}
&\frac{1}{T^2}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_\tau\right] \\
&= \frac{1}{T^2}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_t\right] + \frac{1}{T^2}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{t=1}^{T}\sum_{\tau\neq t}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\varepsilon}_\tau\right] \\
&= \frac{1}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}(\mathcal{M}_4^\varepsilon - 3\sigma_0^4)\sum_{i=1}^{n}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'})_{ii}^2 + \sigma_0^4(\mathrm{tr}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'})^2 + \mathrm{tr}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}) + \mathrm{tr}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_r^0\boldsymbol{\lambda}_s^{0'}))\right] \\
&\quad + \frac{\sigma_0^4(T-1)}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\mathrm{tr}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'})^2\right] \\
&= \frac{1}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}(\mathcal{M}_4^\varepsilon - 3\sigma_0^4)\sum_{i=1}^{n}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'})_{ii}^2 + \sigma_0^4(\mathrm{tr}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}) + \mathrm{tr}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_r^0\boldsymbol{\lambda}_s^{0'}))\right] \\
&\quad + \sigma_0^4\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\mathrm{tr}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'})^2\right]. \qquad \text{(G.8)}
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\frac{1}{T}\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\boldsymbol{\Lambda}^{0'}\boldsymbol{\Lambda}^0\right\|_F^2\right] \\
&= \frac{1}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{i=1}^{n}(\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'})_{ii}^2 + \sigma_0^4(\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_s^0 + \boldsymbol{\lambda}_s^{0'}\boldsymbol{\lambda}_s^0\boldsymbol{\lambda}_r^{0'}\boldsymbol{\lambda}_r^0)\right] \\
&= \frac{(\mathcal{M}_4^\varepsilon - 3\sigma_0^4)}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{i=1}^{n}\lambda_{i,r}^{0^2}\lambda_{i,s}^{0^2}\right] + \frac{\sigma_0^4}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{ir}^0\lambda_{is}^0\lambda_{jr}^0\lambda_{js}^0\right]
\end{aligned}
$$

$$+\ \frac{\sigma_0^4}{T}\mathbb{E}\left[\sum_{r=1}^{R^0}\sum_{s=1}^{R^0}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{is}^{02}\lambda_{jr}^{02}\right]=O\left(\frac{n^2}{T}\right), \tag{G.9}$$

using Assumptions 1 and 2.7. Hence, $||\frac{1}{T}\mathbf{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{\Lambda}^0-\sigma_0^2\mathbf{\Lambda}^{0'}\mathbf{\Lambda}^0||_F=O_P\left(\frac{n}{\sqrt{T}}\right)$. The second part of the lemma follows from analogous steps. □

Proof of Lemma F.1(ix).

$$\begin{aligned}
&\mathbb{E}\left[\sum_{p=1}^{P}\left|\left|\frac{1}{T}\bar{\mathfrak{Z}}_p'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{\Lambda}^0-\sigma_0^2\bar{\mathfrak{Z}}_p'\mathbf{\Lambda}^0\right|\right|_F^2\right]\\
&=\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\left(\frac{1}{T}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_\tau-\sigma_0^2\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\right)^2\right]\\
&=\frac{1}{T^2}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{\tau=1}^{T}\sum_{\tau'=1}^{T}\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_{\tau'}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_{\tau'}\right]-\frac{2\sigma_0^2}{T}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_\tau\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\right]\\
&\quad+\sigma_0^4\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\right]\\
&=\frac{1}{T^2}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{\tau=1}^{T}\sum_{\tau'=1}^{T}\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_{\tau'}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_{\tau'}\right]-\sigma_0^4\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\right].
\end{aligned} \tag{G.10}$$

Using Lemma 3 in Yu et al. (2008),

$$\begin{aligned}
&\frac{1}{T^2}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{\tau=1}^{T}\sum_{\tau'=1}^{T}\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_{\tau'}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_{\tau'}\right]\\
&=\frac{1}{T^2}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_{\tau}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_{\tau}\right]+\frac{1}{T^2}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{\tau=1}^{T}\sum_{\tau'\neq\tau}^{T}\boldsymbol{\varepsilon}_\tau'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_{\tau'}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\varepsilon}_{\tau'}\right]\\
&=\frac{1}{T}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}(\mathcal{M}_4^\varepsilon-3\sigma_0^4)\sum_{i=1}^{n}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}')_{ii}^2+\sigma_0^4(\mathrm{tr}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}')^2+\mathrm{tr}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}')+\mathrm{tr}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\bar{\boldsymbol{Z}}_{pt}\boldsymbol{\lambda}_r^{0'}))\right]\\
&\quad+\frac{\sigma_0^4(T-1)}{T}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\mathrm{tr}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}')^2\right].
\end{aligned} \tag{G.11}$$

Thus,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{p=1}^{P}\left|\left|\frac{1}{T}\bar{\mathcal{Z}}_p'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0-\sigma_0^2\bar{\mathcal{Z}}_p'\boldsymbol{\Lambda}^0\right|\right|_F^2\right]\\
&=\frac{1}{T}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}(\mathcal{M}_4^{\varepsilon}-3\sigma_0^4)\sum_{i=1}^{n}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}')_{ii}^2+\sigma_0^4(\mathrm{tr}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}')+\mathrm{tr}(\boldsymbol{\lambda}_r^0\bar{\boldsymbol{Z}}_{pt}'\bar{\boldsymbol{Z}}_{pt}\boldsymbol{\lambda}_r^{0\prime}))\right]\\
&=\frac{(\mathcal{M}_4^{\varepsilon}-3\sigma_0^4)}{T}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{i=1}^{n}\lambda_{ir}^{0^2}\bar{z}_{pit}^2\right]+\frac{\sigma_0^4}{T}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{ir}^0\bar{z}_{pit}\lambda_{jr}^0\bar{z}_{pjt}\right]\\
&\quad+\frac{\sigma_0^4}{T}\mathbb{E}\left[\sum_{p=1}^{P}\sum_{t=1}^{T}\sum_{r=1}^{R^0}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{ir}^{0^2}\bar{z}_{pjt}^2\right]=O\left(Pn^2\right),
\end{aligned}
\tag{G.12}
$$

which establishes the first part. The second part is obtained similarly. □

Proof of Lemma F.1(x).

$$
\begin{aligned}
\frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) &= \frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon})-\frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})\\
&\quad -\frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})+\frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\\
&=: l_1+...+l_4.
\end{aligned}
\tag{G.13}
$$

Using Lemma 9 in Yu et al. (2008), $l_1=\sigma_0^2+O_P\left(\frac{1}{\sqrt{nT}}\right)$. For l_2,

$$
\frac{1}{nT}|\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon})|\leq\frac{R^0}{nT}||\boldsymbol{\varepsilon}||_2^2=O_P\left(\frac{1}{\min\{n,T\}}\right).
\tag{G.14}
$$

Similarly for l_3 and l_4, which gives the result. □

Proof of Lemma F.2(i). Recall from the discussion of equation (A.1) that $\hat{\boldsymbol{\Lambda}}$ satisfies

$$
\left(\frac{1}{nT}\sum_{t=1}^{T}\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t'\right)\hat{\boldsymbol{\Lambda}}=\hat{\boldsymbol{\Lambda}}\boldsymbol{\Pi},
\tag{G.15}
$$

with the columns of $\hat{\boldsymbol{\Lambda}}$ being R eigenvectors of $\frac{1}{nT}\sum_{t=1}^{T}\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t'$ associated with its R largest eigenvalues, and $\boldsymbol{\Pi}$ being a diagonal $R\times R$ matrix containing the largest R eigenvalues of $\frac{1}{nT}\sum_{t=1}^{T}\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t'$ along its diagonal. With $\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}=\boldsymbol{I}_n+\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q$ by Lemma A.2(i), expand (G.15) as

$$
\hat{\boldsymbol{\Lambda}}\boldsymbol{\Pi}=\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\right)\hat{\boldsymbol{\Lambda}}
$$

$$
\begin{aligned}
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})\boldsymbol{\varepsilon}_t'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\boldsymbol{\varepsilon}_t'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\right)'\right)\hat{\boldsymbol{\Lambda}}
\end{aligned}
$$

$$
\begin{aligned}
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\boldsymbol{\varepsilon}_t'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}))'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&+\left(\frac{1}{nT}\sum_{t=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_t\right)'\right)\hat{\boldsymbol{\Lambda}}\\
&=: \boldsymbol{P}_1+...+\boldsymbol{P}_{25}. \qquad \text{(G.16)}
\end{aligned}
$$

Note that $\boldsymbol{P}_7=\frac{1}{nT}\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}$. Then,

$$\hat{\boldsymbol{\Lambda}}\boldsymbol{\Pi}-\boldsymbol{\Lambda}^0\left(\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\right)=\boldsymbol{P}_1+...+\boldsymbol{P}_6+\boldsymbol{P}_8+...+\boldsymbol{P}_{25}.$$

Since $\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0$ and $\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}$ are both asymptotically invertible, by Assumption 6.3 and Lemma F.2(ii) respectively, let $\boldsymbol{\Sigma}^*:=\text{plim}_{n,T\to\infty}\left(\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\right)^{-1}$ whereby,

$$\hat{\boldsymbol{\Lambda}}\boldsymbol{\Pi}\boldsymbol{\Sigma}^*-\boldsymbol{\Lambda}^0=(\boldsymbol{P}_1+...+\boldsymbol{P}_6+\boldsymbol{P}_8+...+\boldsymbol{P}_{25})\boldsymbol{\Sigma}^*,$$

or

$$\hat{\mathbf{\Lambda}}\boldsymbol{H}^{*-1} - \mathbf{\Lambda}^0 = (\boldsymbol{P}_1 + ... + \boldsymbol{P}_6 + \boldsymbol{P}_8 + ... + \boldsymbol{P}_{25})\mathbf{\Sigma}^*. \tag{G.17}$$

Now,

$$\frac{1}{\sqrt{n}}||\hat{\mathbf{\Lambda}}\boldsymbol{H}^{*-1} - \mathbf{\Lambda}^0||_2 \leq \frac{1}{\sqrt{n}}\left(||\boldsymbol{P}_1||_2 + ... + ||\boldsymbol{P}_6||_2 + ||\boldsymbol{P}_8||_2 + ... + ||\boldsymbol{P}_{25}||_2\right)||\mathbf{\Sigma}^*||_2. \tag{G.18}$$

The probability order of the 24 terms in (G.18) must be examined, though for brevity the calculations for similar terms are omitted. Also note that $||\mathbf{\Sigma}^*||_2 = O_P(1)$. Using Lemmas A.3(i) and F.1(i),

$$\frac{1}{\sqrt{n}}||\boldsymbol{P}_1||_2 \leq \frac{1}{\sqrt{n}}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2||\hat{\mathbf{\Lambda}}||_2 = O_P(||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2). \tag{G.19}$$

Using Lemmas A.2(iii), A.3(i) and F.1(i),

$$\begin{aligned}
\frac{1}{\sqrt{n}}||\boldsymbol{P}_2||_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2||\mathbf{\Lambda}^0\boldsymbol{f}_t^0||_2||\hat{\mathbf{\Lambda}}||_2 \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{\sqrt{nT}}\left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\mathbf{\Lambda}^0\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}}||\hat{\mathbf{\Lambda}}||_2 \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{\sqrt{nT}}\left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}||\mathbf{\Lambda}^0\boldsymbol{F}^{0\prime}||_F||\hat{\mathbf{\Lambda}}||_2 \\
&= O_P(||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2).
\end{aligned} \tag{G.20}$$

Using Lemmas A.3(i) and F.1(i),

$$\begin{aligned}
\frac{1}{\sqrt{n}}||\boldsymbol{P}_3||_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2||\boldsymbol{\varepsilon}_t'\hat{\mathbf{\Lambda}}||_2 \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{\sqrt{nT}}\left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{\varepsilon}_t'\hat{\mathbf{\Lambda}}||_2^2\right)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{n}}\frac{1}{\sqrt{nT}}\left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\varepsilon}'\hat{\mathbf{\Lambda}}||_F \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{\sqrt{nT}}\left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\varepsilon}||_2||\hat{\mathbf{\Lambda}}||_F
\end{aligned}$$

$$= O_P\left(\frac{||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n, T\}}}\right). \tag{G.21}$$

Using Lemmas A.2(i), A.2(iii), A.2(viii), A.3(i) and F.1(i),

$$\begin{aligned}
\frac{1}{\sqrt{n}}||\boldsymbol{P}_4||_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2 \left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0||_2||\hat{\boldsymbol{\Lambda}}||_2 \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{\sqrt{nT}}\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\hat{\boldsymbol{\Lambda}}||_2 \left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}} \\
&= O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2). \qquad (G.22)
\end{aligned}$$

Using Lemmas A.2(i), A.2(vi), A.2(viii), A.3(i) and F.1(i),

$$\begin{aligned}
\frac{1}{\sqrt{n}}||\boldsymbol{P}_5||_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2 \left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\boldsymbol{\varepsilon}_t||_2||\hat{\boldsymbol{\Lambda}}||_2 \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{\sqrt{nT}}\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\hat{\boldsymbol{\Lambda}}||_2 \left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{\varepsilon}_t||_2^2\right)^{\frac{1}{2}} \\
&= O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2). \qquad (G.23)
\end{aligned}$$

Using Lemmas A.2(iii) and F.1(i),

$$\begin{aligned}
\frac{1}{\sqrt{n}}||\boldsymbol{P}_8||_2 &\leq \frac{1}{\sqrt{n}}||\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\boldsymbol{\varepsilon}_t'||_2||\hat{\boldsymbol{\Lambda}}||_2 \leq \frac{1}{\sqrt{n}}\frac{1}{nT}||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F||\boldsymbol{\varepsilon}||_2||\hat{\boldsymbol{\Lambda}}||_2 \\
&= O_P\left(\sqrt{\frac{1}{\min\{n, T\}}}\right). \qquad (G.24)
\end{aligned}$$

Using Lemmas A.2(i), A.2(iii), A.2(viii) and F.1(i),

$$\begin{aligned}
\frac{1}{\sqrt{n}}||\boldsymbol{P}_9||_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0||_2 \left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0||_2||\hat{\boldsymbol{\Lambda}}||_2 \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 \left(\sum_{t=1}^{T}||\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0||_2^2\right)||\hat{\boldsymbol{\Lambda}}||_2 \\
&= \frac{1}{\sqrt{n}}\frac{1}{nT}\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F^2||\hat{\boldsymbol{\Lambda}}||_2
\end{aligned}$$

$$= O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2). \tag{G.25}$$

Using Lemmas A.2(i), A.2(iii), A.2(viii) and F.1(i),

$$\begin{aligned}
\frac{1}{\sqrt{n}}||\boldsymbol{P}_{10}||_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F||\boldsymbol{\varepsilon}||_2\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\hat{\boldsymbol{\Lambda}}||_2 \\
&= O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right).
\end{aligned} \tag{G.26}$$

Using Lemma F.1(i),

$$\frac{1}{\sqrt{n}}||\boldsymbol{P}_{13}||_2 \leq \frac{1}{\sqrt{n}}\frac{1}{nT}||\boldsymbol{\varepsilon}||_2^2||\hat{\boldsymbol{\Lambda}}||_2 = O_P\left(\frac{1}{\min\{n,T\}}\right). \tag{G.27}$$

Using Lemmas A.2(i), A.2(viii) and F.1(i),

$$\frac{1}{\sqrt{n}}||\boldsymbol{P}_{15}||_2 \leq \frac{1}{\sqrt{n}}\frac{1}{nT}||\boldsymbol{\varepsilon}||_2^2\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2 ||\hat{\boldsymbol{\Lambda}}||_2 = O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right). \tag{G.28}$$

Using Lemmas A.2(i), A.2(iii), A.2(viii) and F.1(i),

$$\frac{1}{\sqrt{n}}||\boldsymbol{P}_{19}||_2 = \frac{1}{\sqrt{n}}\frac{1}{nT}\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2^2 ||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F^2||\hat{\boldsymbol{\Lambda}}||_2 = O_P(Q||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2). \tag{G.29}$$

Using Lemmas A.2(i), A.2(iii), A.2(viii) and F.1(i),

$$\frac{1}{\sqrt{n}}||\boldsymbol{P}_{20}||_2 \leq \frac{1}{\sqrt{n}}\frac{1}{nT}\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2^2 ||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F||\boldsymbol{\varepsilon}||_2||\hat{\boldsymbol{\Lambda}}||_2 = O_P\left(\frac{Q||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2}{\sqrt{\min\{n,T\}}}\right). \tag{G.30}$$

Finally, using Lemmas A.2(i), A.2(viii) and F.1(i),

$$\frac{1}{\sqrt{n}}||\boldsymbol{P}_{25}||_2 \leq \frac{1}{\sqrt{n}}\frac{1}{nT}\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2^2 ||\boldsymbol{\varepsilon}||_2^2||\hat{\boldsymbol{\Lambda}}||_2 = O_P\left(\frac{Q||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2}{\min\{n,T\}}\right). \tag{G.31}$$

The orders of the omitted terms follow similarly to those above. Collecting all the terms gives $\frac{1}{\sqrt{n}}||\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2 = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right)$, which establishes the first part of the lemma. For the second part, the first order condition of the maximisation problem (7) yields the condition $\hat{\boldsymbol{F}}' = \frac{1}{n}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{Y} - \sum_{k=1}^{K}\hat{\beta}_k\boldsymbol{\mathcal{X}}_k\right)$. Substituting the true DGP into this expression yields

$$\begin{aligned}\frac{1}{\sqrt{T}}\hat{\boldsymbol{F}}' =& \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{Y} - \sum_{k=1}^{K}\hat{\beta}_k\boldsymbol{\mathcal{X}}_k\right)\\ =& \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\mathfrak{Z}_p\right) + \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\\ &+ \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n\right)\boldsymbol{\varepsilon} + \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} + \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\varepsilon}. \end{aligned}\tag{G.32}$$

Using $\boldsymbol{\Lambda}^0 = \boldsymbol{\Lambda}^0 - \hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1} + \hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}$ and the normalisation $\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}} = \boldsymbol{I}_R$, (G.32) can be rearranged to give

$$\begin{aligned}\frac{1}{\sqrt{T}}(\hat{\boldsymbol{F}}' - \boldsymbol{H}^{*-1}\boldsymbol{F}^{0\prime}) =& \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\mathfrak{Z}_p\right) + \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\\ &+ \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n\right)\boldsymbol{\varepsilon} + \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{\Lambda}^0 - \hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}\right)\boldsymbol{F}^{0\prime} + \frac{1}{n\sqrt{T}}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\varepsilon}\\ =:& \boldsymbol{L}_1 + ... + \boldsymbol{L}_5. \end{aligned}\tag{G.33}$$

Each of these terms is examined. Starting with $\boldsymbol{L}_1$,

$$\begin{aligned}||\boldsymbol{L}_1||_2^2 \le& \frac{1}{n^2T}||\hat{\boldsymbol{\Lambda}}||_2^2\left|\left|\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\mathfrak{Z}_p\right|\right|_F^2\\ =& \frac{1}{n^2T}||\hat{\boldsymbol{\Lambda}}||_2^2\mathrm{tr}\left((\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\boldsymbol{\mathcal{Z}}'\boldsymbol{\mathcal{Z}}(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\right)\\ =& \frac{1}{n}||\hat{\boldsymbol{\Lambda}}||_2^2\mu_1\left(\boldsymbol{\mathcal{H}}_2\right)||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2\\ =& O_P(||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2), \end{aligned}\tag{G.34}$$

using Lemma F.1(i) and Assumption 4.2. Therefore $||\boldsymbol{L}_1||_2 = O_P(||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$. Next,

$$||\boldsymbol{L}_2||_2 \le \frac{1}{n\sqrt{T}}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n||_2||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_2 = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2), \tag{G.35}$$

$$||\boldsymbol{L}_3||_2 \le \frac{1}{n\sqrt{T}}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n||_2||\boldsymbol{\varepsilon}||_2 = O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right), \tag{G.36}$$

$$||\boldsymbol{L}_4||_2 \le \frac{1}{n\sqrt{T}}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Lambda}^0 - \hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{F}^0||_2$$

$$= O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right), \tag{G.37}$$

and

$$||\boldsymbol{L}_5||_2 \leq \frac{1}{n\sqrt{T}}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\varepsilon}||_2 = O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right), \tag{G.38}$$

using Lemmas A.2(iii), A.2(viii) and F.1(ii), and the first part of this lemma whereby $\frac{1}{\sqrt{T}}||\hat{\boldsymbol{F}}' - \boldsymbol{H}^{*-1}\boldsymbol{F}^{0'}||_2 = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right)$. □

Proof of Lemma F.2(ii). Recall that $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}})$ is the maximiser of the penalised average likelihood function $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\Lambda})$. By definition $\mathcal{Q}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}}) \geq \mathcal{Q}(\boldsymbol{\theta}^0, \hat{\boldsymbol{\Lambda}})$, or equivalently,

$$\mathcal{L}(\boldsymbol{\theta}^0, \hat{\boldsymbol{\Lambda}}) - \sum_{p=1}^{P} \varrho_p(\theta_p^0, \gamma_p, \zeta_p) + \sum_{p=1}^{P} \varrho_p(\hat{\theta}_p, \gamma_p, \zeta_p) \leq \mathcal{L}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}}). \tag{G.39}$$

By the same steps as those used to obtain (D.7) in the proof of Proposition 1, it can be shown that

$$\frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\log\left(\sigma_0^2 + o_P(1)\right) + o_P(1) \leq \mathcal{L}(\boldsymbol{\theta}^0, \hat{\boldsymbol{\Lambda}}). \tag{G.40}$$

Now $\mathcal{L}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}})$ can be expanded to give

$$\begin{aligned}
\mathcal{L}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}}) = & \frac{1}{n}\log(\det(\boldsymbol{S}(\hat{\boldsymbol{\rho}}))) - \frac{1}{2}\log\left(\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}))'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\right. \\
& + \frac{2}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}))'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t + \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t \\
& + \frac{2}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \frac{2}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}))'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 \\
& \left. + \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\right) \\
=: & \frac{1}{n}\log(\det(\boldsymbol{S}(\hat{\boldsymbol{\rho}}))) - \frac{1}{2}\log\left(l_1 + ... + l_6\right).
\end{aligned} \tag{G.41}$$

Using similar steps to those for terms k_1,...,k_5 in the proof of Proposition 1, and the result of that Proposition, it can be shown terms l_1 and l_2 are $o_P(1)$, and also that $l_3 = \frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}) + o_P(1)$. Consider the remaining terms. Using Lemma A.2(iii),

$$|l_4| = \frac{2}{nT}\left|\mathrm{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon})'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{F}^{0'})\right|$$

$$
\begin{aligned}
&\leq \frac{2R^0}{nT}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2^2||\boldsymbol{\varepsilon}||_2||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F \\
&= \frac{2R^0}{nT}O_P(\sqrt{\max\{n,T\}})O_P(\sqrt{nT}) = o_P\left(1\right). \qquad \text{(G.42)}
\end{aligned}
$$

Using Lemmas A.2(iii), A.3(i) and Proposition 1,

$$
\begin{aligned}
|l_5| &\leq \frac{2}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})||_2||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2||\boldsymbol{\Lambda}^0||_2||\boldsymbol{f}_t^0||_2 \\
&\leq \frac{2}{\sqrt{nT}}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2||\boldsymbol{\Lambda}^0||_2\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\left(||\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}} \\
&= \frac{2}{\sqrt{nT}}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2||\boldsymbol{\Lambda}^0||_2\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}||\boldsymbol{F}^0||_F \\
&= \frac{2}{\sqrt{nT}}O_P(\sqrt{n})O_P(||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)O_P(\sqrt{T}) = o_P(1). \qquad \text{(G.43)}
\end{aligned}
$$

For term l_6,

$$
\begin{aligned}
l_6 \;&=\; \frac{1}{nT}\sum_{t=1}^{T}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}-\boldsymbol{I}_n)\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}-\boldsymbol{I}_n)\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 \\
&\quad+\frac{2}{nT}\sum_{t=1}^{T}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}-\boldsymbol{I}_n)\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0+\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 \\
&=: l_{6.1}+l_{6.2}+\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0. \qquad \text{(G.44)}
\end{aligned}
$$

By way of Lemmas A.2(iii), A.2(viii) and Proposition 1,

$$
\begin{aligned}
|l_{6.1}| &\leq \frac{1}{nT}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}-\boldsymbol{I}_n||_2^2||\boldsymbol{\Lambda}^0||_2^2\sum_{t=1}^{T}||\boldsymbol{f}_t^0||_2||\boldsymbol{f}_t^0||_2 \\
&\leq \frac{1}{nT}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}-\boldsymbol{I}_n||_2^2||\boldsymbol{\Lambda}^0||_2^2\left(\sum_{t=1}^{T}||\boldsymbol{f}_t^0||_2^2\right) \\
&= \frac{1}{nT}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}-\boldsymbol{I}_n||_2^2||\boldsymbol{\Lambda}^0||_F^2||\boldsymbol{F}^0||_F^2 \\
&= \frac{1}{nT}O_P(Q||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2)O_P(n)O_P(T) = o_P(1). \qquad \text{(G.45)}
\end{aligned}
$$

Similar steps show that $l_{6.2} = o_P(1)$. Returning to (G.40), then

$$
\frac{1}{n}\log(\det(\boldsymbol{S}))-\frac{1}{2}\log\left(\sigma_0^2+o_P(1)\right)+o_P(1)\leq\frac{1}{n}\log(\det(\boldsymbol{S}(\hat{\boldsymbol{\rho}})))
$$

$$-\frac{1}{2}\log\left(\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0+\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})+o_P(1)\right). \tag{G.46}$$

Recall $\frac{1}{n}\log(\det(\boldsymbol{S}))-\frac{1}{2}\log\left(\sigma_0^2\right)-\frac{1}{n}\log(\det(\boldsymbol{S}(\hat{\boldsymbol{\rho}})))=-\frac{1}{2}\log\left(\sigma_0^2\det((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}}\right)$. Using Lemma A.1, $\sigma_0^2\det((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})^{\frac{1}{n}}-\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})'\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})\leq 0$. Therefore multiplying (G.46) by -2, as well as exponentiating, gives

$$\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0+o_P(1)\leq 0. \tag{G.47}$$

The quadratic form $\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0$ is nonnegative and so $\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0=o_P(1)$. Now, $\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0=\frac{1}{nT}\mathrm{tr}(\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0)$. Since the matrix $\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0$ is asymptotically positive definite by Assumption 6.3,

$$\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0=\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\left(\boldsymbol{I}-\frac{1}{n}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Lambda}}'\right)\boldsymbol{\Lambda}^0=\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0-\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0=o_P(1).$$

The matrix $\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0$ is asymptotically invertible by Assumption 6.2 and therefore $\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0$ is also. Since $\det\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\right)=\det\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\right)^2$, $\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}$ converges in probability to an invertible matrix. $\square$

Proof of Lemma F.2(iii). Write

$$\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}=\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*)+\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0\boldsymbol{H}^*=:\boldsymbol{L}_1+\boldsymbol{L}_2. \tag{G.48}$$

By Lemma F.1(ii) $||\boldsymbol{L}_2||_2=O_P\left(\frac{1}{\sqrt{nT}}\right)$. For $\boldsymbol{L}_1$, by decomposition (G.17) in the proof of Lemma F.2(i),

$$\begin{aligned}
\boldsymbol{L}_1&=\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'(\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}-\boldsymbol{\Lambda}^0)\boldsymbol{H}^*\\
&=\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'(\boldsymbol{P}_1+...+\boldsymbol{P}_6+\boldsymbol{P}_7+...+\boldsymbol{P}_{25})\boldsymbol{\Sigma}^*\boldsymbol{H}^*\\
&=\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'(\boldsymbol{P}_2+\boldsymbol{P}_6+\boldsymbol{P}_8+\boldsymbol{P}_9+\boldsymbol{P}_{11}+\boldsymbol{P}_{12}+\boldsymbol{P}_{17})\boldsymbol{\Sigma}^*\boldsymbol{H}^*\\
&\quad+\boldsymbol{O}_P\left(\frac{1}{\min\{n\sqrt{T},T^{1.5}\}}\right)+\boldsymbol{O}_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{\sqrt{nT},T\}}\right)+\boldsymbol{O}_P\left(\frac{Q||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2}{\sqrt{T}}\right)\\
&=:\boldsymbol{L}_{1.1}+...+\boldsymbol{L}_{1.7}+\boldsymbol{O}_P\left(\frac{1}{\min\{n\sqrt{T},T^{1.5}\}}\right)+\boldsymbol{O}_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{\sqrt{nT},T\}}\right)+\boldsymbol{O}_P\left(\frac{Q||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2}{\sqrt{T}}\right).
\end{aligned} \tag{G.49}$$

The probability order of the remaining 7 terms is examined more closely. Starting with $\boldsymbol{L}_{1.1} \coloneqq \frac{1}{nT}\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\boldsymbol{P}_2\boldsymbol{\Sigma}^*\boldsymbol{H}^*$,

$$\begin{aligned}
\boldsymbol{L}_{1.1} &= \frac{1}{nT}\frac{1}{nT}\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\mathcal{Z}_p\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^* \\
&= \frac{1}{nT}\frac{1}{nT}\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\bar{\mathcal{Z}}_p\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^* \\
&\quad + \frac{1}{nT}\frac{1}{nT}\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'(\mathcal{Z}_p - \bar{\mathcal{Z}}_p)\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^* \\
&=: \boldsymbol{L}_{1.1.1} + \boldsymbol{L}_{1.1.2}. \qquad \text{(G.50)}
\end{aligned}$$

Next, using Lemmas A.2(iii), F.1(i) and F.1(vii),

$$\begin{aligned}
||\boldsymbol{L}_{1.1.1}||_2 &\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*||_2\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)^2\right)^{\frac{1}{2}}\left(\sum_{p=1}^{P}||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}} \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^0||_2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*\boldsymbol{H}^*||_2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\left(\sum_{p=1}^{P}||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{nT}}\right), \qquad \text{(G.51)}
\end{aligned}$$

and

$$\begin{aligned}
||\boldsymbol{L}_{1.1.2}||_2 &\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*||_2||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'||_2\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)^2\right)^{\frac{1}{2}}\left(\sum_{p=1}^{P}||\mathcal{Z}_p - \bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}} \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^0||_2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*\boldsymbol{H}^*||_2||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'||_2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\left(\sum_{p=1}^{P}||\mathcal{Z}_p - \bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{\sqrt{nT}, T\}}\right), \qquad \text{(G.52)}
\end{aligned}$$

by Lemmas A.2(iii), F.1(i), F.1(ii) and F.1(v). Hence $||\boldsymbol{L}_{1.1}||_2 = O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{\sqrt{nT}, T\}}\right)$. For $\boldsymbol{L}_{1.2} \coloneqq \frac{1}{nT}\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\boldsymbol{P}_6\boldsymbol{\Sigma}^*\boldsymbol{H}^*$,

$$\boldsymbol{L}_{1.2} = \frac{1}{nT}\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0(\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}))'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*$$

$$= \frac{1}{nT}\frac{1}{nT}\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\mathcal{Z}_p'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*. \tag{G.53}$$

Then,

$$\begin{aligned}
||\boldsymbol{L}_{1.2}||_2 &\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0||_2||\boldsymbol{F}^{0\prime}||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)^2\right)^{\frac{1}{2}}\left(\sum_{p=1}^{P}||\mathcal{Z}_p||_2^2\right)^{\frac{1}{2}} \\
&= \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0||_2||\boldsymbol{F}^{0\prime}||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\left(\sum_{p=1}^{P}||\mathcal{Z}_p||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{nT}}\right),
\end{aligned} \tag{G.54}$$

using Lemmas A.2(iii), A.2(iv), F.1(i) F.1(ii). Next,

$$\boldsymbol{L}_{1.3} \coloneqq \frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{P}_8\boldsymbol{\Sigma}^*\boldsymbol{H}^* = \frac{1}{nT}\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*, \tag{G.55}$$

and

$$||\boldsymbol{L}_{1.3}||_2 \leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0||_2||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2 = O_P\left(\frac{1}{\sqrt{nT}}\right),$$

by Lemmas F.1(i), F.1(ii) and F.1(ii). As for $\boldsymbol{L}_{1.4} \coloneqq \frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{P}_9\boldsymbol{\Sigma}^*\boldsymbol{H}^*$,

$$\boldsymbol{L}_{1.4} = \frac{1}{nT}\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*, \tag{G.56}$$

and therefore

$$\begin{aligned}
||\boldsymbol{L}_{1.4}||_2 &\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0||_2||\boldsymbol{F}^0||_2^2\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2 \\
&= O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{nT}}\right),
\end{aligned} \tag{G.57}$$

using Lemmas A.2(i), A.2(iii), A.2(viii), F.1(i) and F.1(ii). For $\boldsymbol{L}_{1.5} \coloneqq \frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{P}_{11}\boldsymbol{\Sigma}^*\boldsymbol{H}^*$,

$$\boldsymbol{L}_{1.5} = \frac{1}{nT}\frac{1}{nT}\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\mathcal{Z}_p'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*$$

$$
\begin{aligned}
&= \frac{1}{nT}\frac{1}{nT}\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^* \\
&\quad + \frac{1}{nT}\frac{1}{nT}\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}(\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^* \\
&=: \boldsymbol{L}_{1.5.1} + \boldsymbol{L}_{1.5.2}. \qquad \text{(G.58)}
\end{aligned}
$$

By Lemmas F.1(i), F.1(ii) and F.1(vi),

$$
\begin{aligned}
||\boldsymbol{L}_{1.5.1}||_2 &\leq \frac{1}{nT}\frac{1}{nT}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'||_2\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)^2\right)^{\frac{1}{2}}\left(\sum_{p=1}^{P}||\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_2^2\right)^{\frac{1}{2}} \\
&= \frac{1}{nT}\frac{1}{nT}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'||_2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\left(\sum_{p=1}^{P}||\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{T}\right). \qquad \text{(G.59)}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
||\boldsymbol{L}_{1.5.2}||_2 &\leq \frac{1}{nT}\frac{1}{nT}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'||_2||\boldsymbol{\varepsilon}||_2\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)^2\right)^{\frac{1}{2}}\left(\sum_{p=1}^{P}||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2\right)^{\frac{1}{2}} \\
&= \frac{1}{nT}\frac{1}{nT}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'||_2||\boldsymbol{\varepsilon}||_2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\left(\sum_{p=1}^{P}||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n\sqrt{T}, T^{1.5}\}}\right), \qquad \text{(G.60)}
\end{aligned}
$$

using Lemmas F.1(i), F.1(ii) and F.1(v). Hence $||\boldsymbol{L}_{1.5}||_2 = O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{T}\right) + O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n\sqrt{T}, T^{1.5}\}}\right)$. Next for $\boldsymbol{L}_{1.6} \coloneqq \frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{P}_{12}\boldsymbol{\Sigma}^*\boldsymbol{H}^* = \frac{1}{nT}\frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*$,

$$
||\boldsymbol{L}_{1.6}||_2 \leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}||_2^2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2 = O_P\left(\frac{1}{T}\right), \qquad \text{(G.61)}
$$

by Lemmas A.2(iii), F.1(i) and F.1(ii). Finally, for term $\boldsymbol{L}_{1.7} \coloneqq \frac{1}{nT}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{P}_{17}\boldsymbol{\Sigma}^*\boldsymbol{H}^*$

$$
\boldsymbol{L}_{1.7} = \frac{1}{nT}\frac{1}{nT}\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{H}^*. \qquad \text{(G.62)}
$$

Since $\boldsymbol{G}_q$ is UB over q, $\left(\sum_{q=1}^{Q}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{G}_q\boldsymbol{\Lambda}^0||_2^2\right)^{\frac{1}{2}} = O_P(\sqrt{QnT})$ by the same steps as in Lemma F.1(ii). Thus,

$$\begin{aligned}
||\boldsymbol{L}_{1.7}||_2 &\leq \frac{1}{nT}\frac{1}{nT}\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)^2\right)^{\frac{1}{2}}\left(\sum_{q=1}^{Q}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{G}_q\boldsymbol{\Lambda}^0||_2^2\right)^{\frac{1}{2}}||\boldsymbol{F}^0||_2^2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2 \\
&= \frac{1}{nT}\frac{1}{nT}||\boldsymbol{\rho}^0-\hat{\boldsymbol{\rho}}||_2\left(\sum_{q=1}^{Q}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{G}_q\boldsymbol{\Lambda}^0||_2^2\right)^{\frac{1}{2}}||\boldsymbol{F}^0||_2^2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{H}^*||_2 \\
&= O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{nT}}\right) \qquad \text{(G.63)}
\end{aligned}$$

with the additional use of Lemmas A.2(iii) and F.1(i). Combining all the above gives the result

$$\frac{1}{nT}||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}||_2 = O_P\left(\frac{1}{\sqrt{nT}}\right)+O_P\left(\frac{1}{T}\right)+O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2}{\sqrt{T}}\right)+O_P\left(\frac{\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{\sqrt{nT},T\}}\right).$$

□

Proof of Lemma F.2(iv). Decompose

$$-\frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \qquad \text{(G.64)}$$

as

$$\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0-\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 =: \boldsymbol{l}_1+\boldsymbol{l}_2. \qquad \text{(G.65)}$$

Consider the first term

$$\begin{aligned}
\boldsymbol{l}_1 &= \frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\left(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}-\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\right)\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&\quad+\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&=: \boldsymbol{l}_{1.1}+\boldsymbol{l}_{1.2}. \qquad \text{(G.66)}
\end{aligned}$$

For the first of these,

$$||\boldsymbol{l}_{1.1}||_2 \leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}-\boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{\varepsilon}||_2^2||\hat{\boldsymbol{\Lambda}}||_2\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{f}_t^0||_2$$

$$
\begin{aligned}
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{\varepsilon}||_2^2||\hat{\boldsymbol{\Lambda}}||_2 \left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}} ||\boldsymbol{F}^0||_F \\
&= O_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n, T\}}\right) + O_P\left(\frac{\sqrt{P}}{\min\{n^{1.5}, T^{1.5}\}}\right) \qquad \text{(G.67)}
\end{aligned}
$$

using Lemmas A.2(ii), A.2(iv), F.1(i) and F.2(v). For the second term,

$$
\begin{aligned}
\boldsymbol{l}_{1.2} &= \frac{1}{n}\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*)\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&\quad + \frac{1}{n}\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&=: \boldsymbol{l}_{1.2.1} + \boldsymbol{l}_{1.2.2}. \qquad \text{(G.68)}
\end{aligned}
$$

Using Lemmas A.2(ii), A.2(iv), F.2(i) and F.1(ii) and that, by Assumption 6.2, $\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}$ is converging to a fixed positive definite matrix,

$$
\begin{aligned}
||\boldsymbol{l}_{1.2.1}||_2 &\leq \frac{1}{n}\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{\Lambda}^0||_2\left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}||_2||\boldsymbol{\varepsilon}||_2||\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2 \\
&\quad \times||\boldsymbol{\Sigma}^*||_2||\boldsymbol{f}_t^0||_2 \\
&\leq \frac{1}{n}\frac{1}{nT}\frac{1}{nT}||\boldsymbol{\Lambda}^0||_2\left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}||_2||\boldsymbol{\varepsilon}||_2||\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2||\boldsymbol{\Sigma}^*||_2 \\
&\quad \times\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}} ||\boldsymbol{F}^0||_F \\
&= O_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n, \sqrt{nT}\}}\right) + O_P\left(\frac{\sqrt{P}}{\min\{n^{1.5}, \sqrt{nT}\}}\right). \qquad \text{(G.69)}
\end{aligned}
$$

Next,

$$
\begin{aligned}
\boldsymbol{l}_{1.2.2} &= \frac{1}{n^2}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\left(\frac{1}{T}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&\quad + \frac{1}{n^2}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\sigma_0^2\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&=: \boldsymbol{l}_{1.2.2.1} + \boldsymbol{l}_{1.2.2.2}, \qquad \text{(G.70)}
\end{aligned}
$$

where

$$
||\boldsymbol{l}_{1.2.2.1}||_2 \leq \frac{1}{n^2}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{\Lambda}^0||_2\left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2\left|\left|\frac{1}{T}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right|\right|_2
$$

$$
\begin{aligned}
&\times ||\boldsymbol{H}^*||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{f}_t^0||_2 \\
&\leq \frac{1}{n^2}\frac{1}{nT}\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\Lambda}^0||_2\left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2\left|\left|\frac{1}{T}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0-\sigma_0^2\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right|\right|_2 \\
&\times ||\boldsymbol{H}^*||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{F}^0||_2 \\
&= O_P\left(\frac{\sqrt{P}}{n\sqrt{T}}\right) \qquad \text{(G.71)}
\end{aligned}
$$

by Lemmas A.2(iii), A.2(iv) and F.1(viii). Collecting these results together,

$$
\boldsymbol{l}_1 = \boldsymbol{l}_{1.2.2.2} + \boldsymbol{O}_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{P}}{n\sqrt{T}}\right). \qquad \text{(G.72)}
$$

Turning to $\boldsymbol{l}_2$ in (G.65),

$$
\begin{aligned}
\boldsymbol{l}_2 &= -\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*)\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 - \frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&=: \boldsymbol{l}_{2.1} + \boldsymbol{l}_{2.2}, \qquad \text{(G.73)}
\end{aligned}
$$

where

$$
\begin{aligned}
||\boldsymbol{l}_{2.1}||_2 &\leq \frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{\varepsilon}||_2^2||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{f}_t^0||_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\varepsilon}||_2^2||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{F}^0||_F \\
&= O_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + O_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right), \qquad \text{(G.74)}
\end{aligned}
$$

and hence

$$
\boldsymbol{l}_2 = \boldsymbol{l}_{2.2} + \boldsymbol{O}_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right). \qquad \text{(G.75)}
$$

Combining (G.72) and (G.75), ignoring dominated terms, and recalling the definition of $\boldsymbol{P}_{\boldsymbol{\Lambda}^0}$,

$$
\begin{aligned}
\boldsymbol{l}_1 + \boldsymbol{l}_2 = &-\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'-T\sigma_0^2\boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&+\boldsymbol{O}_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{P}}{n\sqrt{T}}\right). \qquad \text{(G.76)}
\end{aligned}
$$

Now,

$$
\begin{aligned}
&-\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'-T\sigma_0^2\boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\\
=&-\frac{1}{nT}\frac{1}{n}\sum_{t=1}^{T}\bar{\boldsymbol{Z}}_t'\left(\frac{1}{T}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'-\sigma_0^2\boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0+\frac{1}{nT}\frac{1}{n}\sum_{t=1}^{T}(\boldsymbol{Z}_t-\bar{\boldsymbol{Z}}_t)'\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\\
&-\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t-\bar{\boldsymbol{Z}}_t)'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\\
=:&\ \boldsymbol{j}_1+\boldsymbol{j}_2+\boldsymbol{j}_3. \qquad (G.77)
\end{aligned}
$$

Consider each of these three terms. Using Lemmas A.2(ii) and F.1(ix),

$$
\begin{aligned}
||\boldsymbol{j}_1||_2^2 &= \frac{1}{n^4T^2}\sum_{p=1}^{P}\operatorname{tr}\left(\bar{\boldsymbol{\mathcal{Z}}}_p'\left(\frac{1}{T}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'-\sigma_0^2\boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\boldsymbol{H}^*\boldsymbol{\Sigma}^*\boldsymbol{F}^{0'}\right)^2\\
&\le \frac{1}{n^4T^2}R^2||\boldsymbol{H}^*||_2^2||\boldsymbol{\Sigma}^*||_2^2||\boldsymbol{F}^{0'}||_2^2\sum_{p=1}^{P}\left|\left|\bar{\boldsymbol{\mathcal{Z}}}_p'\left(\frac{1}{T}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'-\sigma_0^2\boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\right|\right|_2^2\\
&= O_P\left(\frac{P}{n^2T}\right), \qquad (G.78)
\end{aligned}
$$

whereby $||\boldsymbol{j}_1||_2 = O_P\left(\frac{\sqrt{P}}{n\sqrt{T}}\right)$. By similar steps, and using Lemmas A.2(ii) and F.1(v), it can be shown that $||\boldsymbol{j}_2||_2$ and $||\boldsymbol{j}_3||_2$ are $O_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},\sqrt{n}T\}}\right)$. Combining all these results and ignoring dominated terms gives

$$
-\frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0=\boldsymbol{O}_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right)+\boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right). \qquad (G.79)
$$

□

Proof of Lemma F.2(v). Following Lemma A.7(i) in Bai (2009), first note that

$$
\begin{aligned}
\left|\left|\frac{1}{n}\boldsymbol{\Lambda}^{0'}(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*)\right|\right|_2 &\le \frac{1}{n}||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2\\
&= O_P(\sqrt{Q}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right), \qquad (G.80)
\end{aligned}
$$

and

$$
\left|\left|\frac{1}{n}\hat{\boldsymbol{\Lambda}}'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*)\right|\right|_2 \le \frac{1}{n}||\hat{\boldsymbol{\Lambda}}||_2||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2
$$

$$= O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right). \tag{G.81}$$

Thus,

$$\begin{aligned}\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*) &= \frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}} - \frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}^* \\ &= \boldsymbol{O}_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right),\end{aligned} \tag{G.82}$$

and

$$\begin{aligned}\frac{1}{n}\hat{\boldsymbol{\Lambda}}'(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*) &= \frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}} - \frac{1}{n}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{H}^* \\ &= \boldsymbol{I}_R - \frac{1}{n}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{H}^* \\ &= \boldsymbol{O}_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right).\end{aligned} \tag{G.83}$$

Left multiply (G.82) by $\boldsymbol{H}^{*\prime}$ and use the transpose of (G.83) to obtain

$$\boldsymbol{I}_R - \frac{1}{n}\boldsymbol{H}^{*\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}^* = \boldsymbol{O}_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right). \tag{G.84}$$

Next, noting that $\boldsymbol{P}_{\boldsymbol{\Lambda}^0\boldsymbol{H}} = \boldsymbol{P}_{\boldsymbol{\Lambda}^0}$ for any invertible $\boldsymbol{H}$,

$$\begin{aligned}\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0} &= \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0\boldsymbol{H}} \\ &= (\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}})^{-1}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})' + (\hat{\boldsymbol{\Lambda}} - \boldsymbol{H}\boldsymbol{\Lambda}^0)(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}})^{-1}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime} \\ &\quad + \boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}})^{-1}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})' + \boldsymbol{\Lambda}^0\boldsymbol{H}((\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}})^{-1} - (\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H})^{-1})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime} \\ &= \frac{1}{n}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})\left(\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}\right)^{-1}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})' + \frac{1}{n}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})\left(\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}\right)^{-1}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime} \\ &\quad + \frac{1}{n}\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}\right)^{-1}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{H}\boldsymbol{\Lambda}^0)' \\ &\quad + \frac{1}{n}\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\left(\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}\right)^{-1} - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}.\end{aligned} \tag{G.85}$$

Recalling $\frac{1}{n}\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}} = \boldsymbol{I}_R$, (G.85) becomes

$$\begin{aligned}\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0} \;=\; &\frac{1}{n}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})' + \frac{1}{n}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime} \\ &+ \frac{1}{n}\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})' + \frac{1}{n}\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\end{aligned}$$

$$=: \boldsymbol{L}_1 + \boldsymbol{L}_2 + \boldsymbol{L}_3 + \boldsymbol{L}_4.$$

Choosing $\boldsymbol{H}$ to be the $\boldsymbol{H}^*$ given in Lemma F.2(i), Lemmas A.2(iii), F.2(i) and equation (G.84) can be exploited to give

$$||\boldsymbol{L}_1||_2 \leq \frac{1}{n}||\boldsymbol{\Lambda} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2^2 = O_P(Q||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2) + O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + O_P\left(\frac{1}{\min\{n,T\}}\right)$$

$$||\boldsymbol{L}_2||_2 = ||\boldsymbol{L}_3||_2 \leq \frac{1}{\sqrt{n}}||\boldsymbol{\Lambda} - \boldsymbol{\Lambda}^0\boldsymbol{H}^*||_2\frac{1}{\sqrt{n}}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}^*||_2 = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right)$$

$$||\boldsymbol{L}_4||_2 \leq \frac{1}{n}||\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{H}^*||_2^2\left|\left|\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}^{*\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}^*\right)^{-1}\right|\right|_2 = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right).$$

The three results above gives

$$||\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2 = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right), \tag{G.86}$$

which concludes the first part of the lemma. The second part can be shown similarly, using Lemmas F.2(i) and F.2(vi). □

Proof of Lemma F.2(vi). From equation (G.33) in the proof of Lemma F.2(i),

$$\begin{aligned}\frac{1}{\sqrt{T}}\hat{\boldsymbol{F}}' &= \frac{1}{\sqrt{T}n}\hat{\boldsymbol{\Lambda}}'\left(\sum_{p=1}^{P}(\theta_p^0 - \hat{\theta}_p)\mathcal{Z}_p\right) + \frac{1}{\sqrt{T}n}\hat{\boldsymbol{\Lambda}}'\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n\right)\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} \\ &\quad + \frac{1}{\sqrt{T}n}\hat{\boldsymbol{\Lambda}}\left(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1} - \boldsymbol{I}_n\right)\boldsymbol{\varepsilon} + \frac{1}{\sqrt{T}n}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} + \frac{1}{\sqrt{T}n}\hat{\boldsymbol{\Lambda}}\boldsymbol{\varepsilon} \\ &=: \boldsymbol{L}_1 + \boldsymbol{L}_2 + \boldsymbol{L}_3 + \boldsymbol{L}_4 + \boldsymbol{L}_5.\end{aligned} \tag{G.87}$$

It was shown in the proof of Lemma F.2(i) that $||\boldsymbol{L}_1||_2, ||\boldsymbol{L}_2||_2, ||\boldsymbol{L}_3||_2, ||\boldsymbol{L}_5||_2 = O_P(1)$. For $\boldsymbol{L}_4$, $||\boldsymbol{L}_4^*||_2 \leq \frac{1}{\sqrt{T}n}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_2 = O_P(1)$, using Lemmas A.2(iii) and F.1(i), which gives the result. □

Proof of Lemma F.2(vii).

$$\begin{aligned}&\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon}) \\ &= \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}(\mathcal{Z}_1'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0} - \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots \\ \mathrm{tr}(\mathcal{Z}_P'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0} - \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\end{pmatrix} + \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}((\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0} - \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots \\ \mathrm{tr}((\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0} - \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\end{aligned}$$

$$
\begin{aligned}
&+\ \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}-\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}-\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}+\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*'}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon})\\
&=: \boldsymbol{T}_1+\boldsymbol{T}_2+\boldsymbol{T}_3+\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*'}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}). \qquad (G.88)
\end{aligned}
$$

Terms $\boldsymbol{T}_1, \boldsymbol{T}_2$ and $\boldsymbol{T}_3$ are now examined. Starting with $\boldsymbol{T}_1$,

$$
\begin{aligned}
\boldsymbol{T}_1 &= \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\mathcal{Z}_1'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}(\mathcal{Z}_P'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\end{pmatrix}\\
&+\frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\mathcal{Z}_1'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}(\mathcal{Z}_P'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\end{pmatrix}\\
&+\frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\mathcal{Z}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}(\mathcal{Z}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\end{pmatrix}\\
&+\frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}\left(\mathcal{Z}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R-\left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0'}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots \\ \mathrm{tr}\left(\mathcal{Z}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R-\left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0'}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\end{pmatrix}\\
&=: \boldsymbol{T}_{1.1}+\boldsymbol{T}_{1.2}+\boldsymbol{T}_{1.3}+\boldsymbol{T}_{1.4}. \qquad (G.89)
\end{aligned}
$$

For term $\boldsymbol{T}_{1.1}$,

$$
\begin{aligned}
||\boldsymbol{T}_{1.1}||_2 &\le \frac{1}{\sqrt{nT}}\frac{R}{n}\left(\sum_{p=1}^{P}||\mathcal{Z}_p||_2^2\right)^{\frac{1}{2}}||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}||_2^2||\boldsymbol{\varepsilon}||_2||\boldsymbol{M}_{\boldsymbol{F}^0}||_2\\
&= O_P\left(\sqrt{P}Q\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2\right)+O_P\left(\frac{\sqrt{Q}\sqrt{P}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right)\\
&\quad +O_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{n,T\}}\right), \qquad (G.90)
\end{aligned}
$$

using Lemmas A.2(iv) and F.2(i). Next,

$$\begin{aligned}
\boldsymbol{T}_{1.2} = \; & \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\bar{\mathcal{Z}}_1'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon})\\ \vdots \\ \mathrm{tr}(\bar{\mathcal{Z}}_P'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon})\end{pmatrix} + \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}((\mathcal{Z}_1-\bar{\mathcal{Z}}_1)'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon})\\ \vdots \\ \mathrm{tr}((\mathcal{Z}_P-\bar{\mathcal{Z}}_P)'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon})\end{pmatrix}\\
& - \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\bar{\mathcal{Z}}_1'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}(\bar{\mathcal{Z}}_P'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\end{pmatrix}\\
& - \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}((\mathcal{Z}_1-\bar{\mathcal{Z}}_1)'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}((\mathcal{Z}_P-\bar{\mathcal{Z}}_P)'(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H})\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\end{pmatrix}\\
=: \; & \boldsymbol{T}_{1.2.1}+\boldsymbol{T}_{1.2.2}+\boldsymbol{T}_{1.2.3}+\boldsymbol{T}_{1.2.4}. \qquad \text{(G.91)}
\end{aligned}$$

Considering the four terms above,

$$\begin{aligned}
||\boldsymbol{T}_{1.2.1}||_2 &\leq \frac{1}{\sqrt{nT}}\frac{R}{n}||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}||_2||\boldsymbol{H}||_2\left(\sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_p'||_2^2\right)^{\frac{1}{2}}\\
&= O_P\left(\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\right)+O_P\left(\sqrt{\frac{P}{\min\{n,T\}}}\right), \qquad \text{(G.92)}
\end{aligned}$$

$$\begin{aligned}
||\boldsymbol{T}_{1.2.2}||_2 &\leq \frac{1}{\sqrt{nT}}\frac{R}{n}||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}||_2||\boldsymbol{H}||_2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}||_2\left(\sum_{p=1}^{P}||\mathcal{Z}_p-\bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}}\\
&= O_P\left(\frac{\sqrt{QP}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{n}}\right)+O_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{n,\sqrt{nT}\}}\right), \qquad \text{(G.93)}
\end{aligned}$$

$$\begin{aligned}
||\boldsymbol{T}_{1.2.3}||_2 &\leq \frac{1}{\sqrt{nT}}\frac{R}{nT}||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}||_2||\boldsymbol{H}||_2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2\left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2||\boldsymbol{F}^0||_2\left(\sum_{p=1}^{P}||\bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}}\\
&= O_P\left(\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\right)+O_P\left(\sqrt{\frac{P}{\min\{n,T\}}}\right), \qquad \text{(G.94)}
\end{aligned}$$

and

$$||\boldsymbol{T}_{1.2.4}||_2 \leq \frac{1}{\sqrt{nT}}\frac{R}{nT}||\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\boldsymbol{H}||_2||\boldsymbol{H}||_2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2\left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2||\boldsymbol{F}^0||_2\left(\sum_{p=1}^{P}||\mathcal{Z}_p-\bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}}$$

$$= O_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right) + O_P\left(\frac{\sqrt{P}}{\min\{n,T\}}\right), \tag{G.95}$$

using Lemmas A.2(iii), A.2(iv), F.1(ii), F.1(v), F.1(vii) and F.2(i). For term $\boldsymbol{T}_{1.3}$,

$$\begin{aligned}
\boldsymbol{T}_{1.3} &= \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\bar{\mathfrak{Z}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon})\\ \vdots \\ \mathrm{tr}(\bar{\mathfrak{Z}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon})\end{pmatrix} + \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon})\\ \vdots \\ \mathrm{tr}((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon})\end{pmatrix}\\
&- \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\bar{\mathfrak{Z}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}(\bar{\mathfrak{Z}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\end{pmatrix}\\
&- \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\\ \vdots \\ \mathrm{tr}((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)'\boldsymbol{\Lambda}^0\boldsymbol{H}(\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\end{pmatrix}\\
&=: \boldsymbol{T}_{1.3.1} + \boldsymbol{T}_{1.3.2} + \boldsymbol{T}_{1.3.3} + \boldsymbol{T}_{1.3.4}.
\end{aligned} \tag{G.96}$$

Consider terms $\boldsymbol{T}_{1.3.2}$ and $\boldsymbol{T}_{1.3.4}$. One has

$$\begin{aligned}
||\boldsymbol{T}_{1.3.2}||_2 &\le \frac{1}{\sqrt{nT}}\frac{R}{n}||\boldsymbol{H}||_2||\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}||_2||\boldsymbol{\varepsilon}||_2\left(\sum_{p=1}^{P}||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\Lambda}^0||_2\\
&= O_P\left(\frac{\sqrt{QP}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right) + O_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{n,T\}}\right),
\end{aligned} \tag{G.97}$$

and

$$\begin{aligned}
||\boldsymbol{T}_{1.3.4}||_2 &\le \frac{1}{\sqrt{nT}}\frac{R}{nT}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2||\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\boldsymbol{H}||_2^2||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2\left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2||\boldsymbol{F}^0||_2\left(\sum_{P=1}^{P}||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2\right)^{\frac{1}{2}}\\
&= O_P\left(\frac{\sqrt{QP}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{T}}\right) + O_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{\sqrt{nT},T\}}\right),
\end{aligned} \tag{G.98}$$

using Lemmas A.2(iii), F.1(ii), F.1(v) and F.2(i). The analysis of term $\boldsymbol{T}_{1.3.1}$ is more involved. Using the same expansion as in the proof of Lemma F.2(iii) one arrives at

$$\boldsymbol{T}_{1.3.1} = \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix}\mathrm{tr}(\bar{\mathfrak{Z}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'(\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{-1} - \boldsymbol{\Lambda}^0)'\boldsymbol{\varepsilon})\\ \vdots \\ \mathrm{tr}(\bar{\mathfrak{Z}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'(\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{-1} - \boldsymbol{\Lambda}^0)'\boldsymbol{\varepsilon})\end{pmatrix}$$

$$= \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix} \mathrm{tr}(\bar{\mathfrak{Z}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'(\boldsymbol{P}_2+\boldsymbol{P}_6+\boldsymbol{P}_8+\boldsymbol{P}_9+\boldsymbol{P}_{11}+\boldsymbol{P}_{12}+\boldsymbol{P}_{17})'\boldsymbol{\varepsilon}) \\ \vdots \\ \mathrm{tr}(\bar{\mathfrak{Z}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'(\boldsymbol{P}_2+\boldsymbol{P}_6+\boldsymbol{P}_8+\boldsymbol{P}_9+\boldsymbol{P}_{11}+\boldsymbol{P}_{12}+\boldsymbol{P}_{17})'\boldsymbol{\varepsilon}) \end{pmatrix}$$
$$+\boldsymbol{O}_P\left(\frac{\sqrt{QP}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right)+\boldsymbol{O}_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{n,T\}}\right)$$
$$=: \boldsymbol{T}_{1.3.1.1}+\ldots+\boldsymbol{T}_{1.3.1.7}+\boldsymbol{O}_P\left(Q\sqrt{P}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2\right)$$
$$+\boldsymbol{O}_P\left(\frac{\sqrt{QP}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right)+\boldsymbol{O}_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{n,T\}}\right). \tag{G.99}$$

Each of the remaining 7 subterms in (G.99) must also be considered. Using Lemmas A.2(iii), F.1(i) and F.1(vii),

$$||\boldsymbol{T}_{1.3.1.1}||_2 \le \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{R}{nT}||\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{H}||_2^2||\boldsymbol{F}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\left(\sum_{p=1}^{P}\sum_{p'=1}^{P}||\bar{\mathfrak{Z}}_{p'}'\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_2^2\right)^{\frac{1}{2}}$$
$$= O_P\left(P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\right). \tag{G.100}$$

Using Lemmas A.2(iii), A.3(i), F.1(i) and F.1(vii),

$$||\boldsymbol{T}_{1.3.1.2}||_2 \le \frac{1}{n}\frac{R}{nT}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2^2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{F}^0||_F\left(\sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_2^2\right)^{\frac{1}{2}}\left(\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}$$
$$= O_P(\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2). \tag{G.101}$$

Using Lemmas A.2(iii), F.1(i), F.1(ii) F.1(vii),

$$||\boldsymbol{T}_{1.3.1.3}||_2 \le \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{R}{nT}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2^2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2\left(\sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p||_2^2\right)^{\frac{1}{2}}$$
$$= O_P\left(\sqrt{\frac{P}{T}}\right). \tag{G.102}$$

Also, using Lemmas A.2(i), A.2(iii), A.2(viii), F.1(i) and F.1(vii),

$$||\boldsymbol{T}_{1.3.1.4}||_2 \le \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{R}{nT}||\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{H}||_2^2||\hat{\boldsymbol{\Lambda}}||_2\left|\left|\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2||\boldsymbol{F}^0||_2^2\left(\sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\bar{\mathfrak{Z}}_p'||_2^2\right)^{\frac{1}{2}}$$
$$= O_P(\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2). \tag{G.103}$$

For term $\boldsymbol{T}_{1.3.1.5}$,

$$\boldsymbol{T}_{1.3.1.5} = \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\sum_{p'=1}^{P}(\theta_{p'}^0 - \hat{\theta}_{p'})\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\bar{\mathcal{Z}}_{p'}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_1'\right) \\ \vdots \\ \operatorname{tr}\left(\sum_{p'=1}^{P}(\theta_{p'}^0 - \hat{\theta}_{p'})\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\bar{\mathcal{Z}}_{p'}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_P'\right) \end{pmatrix}$$
$$+ \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\sum_{p'=1}^{P}(\theta_{p'}^0 - \hat{\theta}_{p'})\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'(\mathcal{Z}_{p'} - \bar{\mathcal{Z}}_{p'})\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_1'\right) \\ \vdots \\ \operatorname{tr}\left(\sum_{p'=1}^{P}(\theta_{p'}^0 - \hat{\theta}_{p'})\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'(\mathcal{Z}_{p'} - \bar{\mathcal{Z}}_{p'})\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_P'\right) \end{pmatrix}$$
$$=: \boldsymbol{T}_{1.3.1.5.1} + \boldsymbol{T}_{1.3.1.5.2}, \tag{G.104}$$

where

$$||\boldsymbol{T}_{1.3.1.5.1}||^2 \le \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{1}{nT}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\left(\sum_{p'=1}^{P}||\bar{\mathcal{Z}}_{p'}\boldsymbol{\varepsilon}'||_2^2\right)^{\frac{1}{2}}\left(\sum_{p=1}^{P}||\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_p'||_2^2\right)^{\frac{1}{2}}$$
$$= O_P\left(\frac{P\sqrt{n}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{T}}\right), \tag{G.105}$$

and

$$||\boldsymbol{T}_{1.3.1.5.2}||_2 \le \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{R}{nT}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2^2||\hat{\boldsymbol{\Lambda}}||_2|||\boldsymbol{\varepsilon}||_2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\left(\sum_{p'=1}^{P}||\mathcal{Z}_{p'} - \bar{\mathcal{Z}}_{p'}||_2^2\right)^{\frac{1}{2}}\left(\sum_{p=1}^{P}||\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_p'||_2^2\right)^{\frac{1}{2}}$$
$$= O_P\left(\frac{P\sqrt{n}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n, T\}}\right), \tag{G.106}$$

using Lemmas A.2(iii), F.1(i), F.1(v) and F.1(vi). For term $\boldsymbol{T}_{1.3.1.6}$,

$$\boldsymbol{T}_{1.3.1.6} = \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}(\bar{\mathcal{Z}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}) \\ \vdots \\ \operatorname{tr}(\bar{\mathcal{Z}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}) \end{pmatrix}. \tag{G.107}$$

Lastly, using Lemmas A.2(iii), F.1(i) and F.1(vii),

$$||\boldsymbol{T}_{1.3.1.7}||_2 \le \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{R}{nT}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2^2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{F}^0||_2^2||\boldsymbol{\rho}^0 - \hat{\boldsymbol{\rho}}||_2\left(\sum_{p=1}^{P}\sum_{q=1}^{Q}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{G}_q'\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_p'||_2^2\right)^{\frac{1}{2}}$$
$$= O_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{T}}\right). \tag{G.108}$$

Combining all the above results gives

$$
\begin{aligned}
\boldsymbol{T}_{1.3.1} = \boldsymbol{T}_{1.3.1.6} + \ & \boldsymbol{O}_P\left(\sqrt{\frac{P}{T}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{n,T\}}\right) + \boldsymbol{O}_P\left(\frac{P\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) \\
& + \boldsymbol{O}_P\left(P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\right) + \boldsymbol{O}_P\left(Q\sqrt{P}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2\right) \\
= \ & \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{1}{nT}\boldsymbol{T}_{1.3.1.6} + \boldsymbol{o}_P(1),
\end{aligned} \tag{G.109}
$$

since $T/n \to c$ by Assumption 6.4, and $||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2$ is at least of order a_{nT} by Proposition 1. Analogous steps for term $\boldsymbol{T}_{1.3.3}$ yield

$$
\boldsymbol{T}_{1.3.3} = \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}(\bar{\boldsymbol{\mathfrak{Z}}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \\ \vdots \\ \operatorname{tr}(\bar{\boldsymbol{\mathfrak{Z}}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \end{pmatrix} + \boldsymbol{o}_P(1). \tag{G.110}
$$

Together (G.109) and (G.110) give the result

$$
\begin{aligned}
\boldsymbol{T}_{1.3} = \boldsymbol{T}_{1.3.1.6} & - \frac{1}{\sqrt{nT}}\frac{1}{n}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}(\bar{\boldsymbol{\mathfrak{Z}}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \\ \vdots \\ \operatorname{tr}(\bar{\boldsymbol{\mathfrak{Z}}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \end{pmatrix} + \boldsymbol{o}_P(1) \\
= \ & \frac{1}{\sqrt{nT}}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\boldsymbol{\mathfrak{Z}}}_1' - \sigma_0^2\boldsymbol{F}^{0\prime}\bar{\boldsymbol{\mathfrak{Z}}}_1'\right)\right) \\ \vdots \\ \operatorname{tr}\left(\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\boldsymbol{\mathfrak{Z}}}_1' - \sigma_0^2\boldsymbol{F}^{0\prime}\bar{\boldsymbol{\mathfrak{Z}}}_P'\right)\right) \end{pmatrix} \\
& + \frac{1}{\sqrt{nT}}\frac{\sigma_0^2}{nT}\begin{pmatrix} \operatorname{tr}\left(\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\bar{\boldsymbol{\mathfrak{Z}}}_1'\right) \\ \vdots \\ \operatorname{tr}\left(\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\bar{\boldsymbol{\mathfrak{Z}}}_p'\right) \end{pmatrix} \\
& - \frac{1}{\sqrt{nT}}\frac{1}{nT}\frac{1}{T}\begin{pmatrix} \operatorname{tr}\left(\bar{\boldsymbol{\mathfrak{Z}}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0 - \sigma_0^2\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0\prime}\right) \\ \vdots \\ \operatorname{tr}\left(\bar{\boldsymbol{\mathfrak{Z}}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0 - \sigma_0^2\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0\prime}\right) \end{pmatrix} \\
& - \frac{1}{\sqrt{nT}}\frac{\sigma_0^2}{nT}\begin{pmatrix} \operatorname{tr}\left(\bar{\boldsymbol{\mathfrak{Z}}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{P}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \operatorname{tr}(\bar{\boldsymbol{\mathfrak{Z}}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{P}_{\boldsymbol{F}^0}) \end{pmatrix} + \boldsymbol{o}_P(1)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{nT}}\frac{1}{nT}\begin{pmatrix} \operatorname{tr}\left(\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\boldsymbol{\mathcal{Z}}}_1' - \sigma_0^2\boldsymbol{F}^{0\prime}\bar{\boldsymbol{\mathcal{Z}}}_1'\right)\right) \\ \vdots \\ \operatorname{tr}\left(\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\boldsymbol{\mathcal{Z}}}_1' - \sigma_0^2\boldsymbol{F}^{0\prime}\bar{\boldsymbol{\mathcal{Z}}}_P'\right)\right) \end{pmatrix} \\
&\quad - \frac{1}{\sqrt{nT}}\frac{1}{nT}\frac{1}{T}\begin{pmatrix} \operatorname{tr}\left(\bar{\boldsymbol{\mathcal{Z}}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0 - \sigma_0^2\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0\prime}\right) \\ \vdots \\ \operatorname{tr}\left(\bar{\boldsymbol{\mathcal{Z}}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{H}'\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0 - \sigma_0^2\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0\prime}\right) \end{pmatrix} \\
&\quad + \boldsymbol{o}_P(1) \\
&=: \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{o}_P(1). \qquad \text{(G.111)}
\end{aligned}
$$

For terms $\boldsymbol{a}$ and $\boldsymbol{b}$,

$$
\begin{aligned}
||\boldsymbol{a}||_2 &\leq \frac{1}{\sqrt{nT}}\frac{1}{nT}||\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{H}||_2^2||\hat{\boldsymbol{\Lambda}}||_2\left(\sum_{p=1}^{P}\left|\left|\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\bar{\boldsymbol{\mathcal{Z}}}_p' - \sigma_0^2\boldsymbol{F}^{0\prime}\bar{\boldsymbol{\mathcal{Z}}}_p'\right|\right|_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\sqrt{\frac{P}{T}}\right), \qquad \text{(G.112)}
\end{aligned}
$$

and

$$
\begin{aligned}
||\boldsymbol{b}||_2 \leq{}& \frac{1}{\sqrt{nT}}\frac{1}{nT}\frac{1}{T}||\boldsymbol{\Lambda}^0||_2^2||\boldsymbol{H}||_2^2||\hat{\boldsymbol{\Lambda}}||_2\left|\left|\frac{1}{n}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\boldsymbol{F}^0 - \sigma_0^2\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right|\right|_2\left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2 \\
&\times ||\boldsymbol{F}^0||_2\left(\sum_{p=1}^{P}||\bar{\boldsymbol{\mathcal{Z}}}_p||_2^2\right)^{\frac{1}{2}} = O_P\left(\sqrt{\frac{P}{T}}\right), \qquad \text{(G.113)}
\end{aligned}
$$

using Lemmas A.2(iii), A.2(iv), F.1(i), F.1(viii) and F.1(ix). Therefore $||\boldsymbol{T}_{1.3}||_2 = o_P(1)$. Next,

$$
\begin{aligned}
\boldsymbol{T}_{1.4} ={}& \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix} \operatorname{tr}\left(\bar{\boldsymbol{\mathcal{Z}}}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\right) \\ \vdots \\ \operatorname{tr}\left(\bar{\boldsymbol{\mathcal{Z}}}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\right) \end{pmatrix} \\
&+ \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix} \operatorname{tr}\left((\boldsymbol{\mathcal{Z}}_1 - \bar{\boldsymbol{\mathcal{Z}}}_1)'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\right) \\ \vdots \\ \operatorname{tr}\left((\boldsymbol{\mathcal{Z}}_P - \bar{\boldsymbol{\mathcal{Z}}}_P)'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\right) \end{pmatrix}
\end{aligned}
$$

$$
- \frac{1}{\sqrt{nT}}\frac{1}{n}\begin{pmatrix} \mathrm{tr}\left(\mathcal{Z}_1'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left(\mathcal{Z}_P'\boldsymbol{\Lambda}^0\boldsymbol{H}\left(\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right)\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right) \end{pmatrix}
$$

$$
=: \boldsymbol{T}_{1.4.1} + \boldsymbol{T}_{1.4.2} + \boldsymbol{T}_{1.4.3}. \tag{G.114}
$$

First,

$$
\begin{aligned}
||\boldsymbol{T}_{1.4.1}||_2 &\leq \frac{1}{\sqrt{nT}}\frac{R}{n}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2^2\left\|\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right\|_2\left(\sum_{p=1}^{P}||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\bar{\mathcal{Z}}_p'||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\sqrt{QP}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\right) + O_P\left(\sqrt{\frac{P}{\min\{n,T\}}}\right).
\end{aligned} \tag{G.115}
$$

Second,

$$
\begin{aligned}
||\boldsymbol{T}_{1.4.2}||_2 &\leq \frac{1}{\sqrt{nT}}\frac{R}{n}\left(\sum_{p=1}^{P}||\mathcal{Z}_p - \bar{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2^2\left\|\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right\|_2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}||_2 \\
&= O_P\left(\frac{\sqrt{QP}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{n}}\right) + O_P\left(\frac{\sqrt{P}\sqrt{\max\{n,T\}}}{\min\{n,\sqrt{nT}\}}\right).
\end{aligned} \tag{G.116}
$$

Third,

$$
\begin{aligned}
||\boldsymbol{T}_{1.4.3}||_2 &\leq \frac{R}{\sqrt{nT}}\frac{1}{nT}\left(\sum_{p=1}^{P}||\mathcal{Z}_p||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\Lambda}^0||_2||\boldsymbol{H}||_2^2\left\|\boldsymbol{I}_R - \left(\frac{1}{n}\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\right)^{-1}\right\|_2||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2 \\
&\quad \times \left\|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\right\|_2||\boldsymbol{F}^0||_2 \\
&= O_P\left(\sqrt{QP}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2\right) + O_P\left(\sqrt{\frac{P}{\min\{n,T\}}}\right).
\end{aligned} \tag{G.117}
$$

This gives the result $\boldsymbol{T}_1 = \boldsymbol{o}_P(1)$ under Assumption 6.4 and Proposition 1. Next, note that, because $\boldsymbol{M}_{\boldsymbol{F}^0}\boldsymbol{F}^0 = \boldsymbol{0}_{T\times R^0}$, $\boldsymbol{T}_2 = \boldsymbol{0}_{P\times 1}$. Thus, it remains only to examine $\boldsymbol{T}_3$ where

$$
\begin{aligned}
||\boldsymbol{T}_3||_2 &\leq \frac{1}{\sqrt{nT}}\left(\sum_{q=1}^{Q}||\boldsymbol{G}_q||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\varepsilon}||_2^2||\boldsymbol{P}_{\boldsymbol{\Lambda}^0} - \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{M}_{\boldsymbol{F}^0}||_2 \\
&= O_P\left(\frac{\sqrt{Q}\sqrt{\max\{n,T\}}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{\min\{n,T\}}}\right) + O_P\left(\frac{\sqrt{\max\{n,T\}}}{\min\{n,T\}}\right).
\end{aligned} \tag{G.118}
$$

Collecting all the terms above, and, with $T/n \to c$ by Assumption 6.4,

$$\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*'}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon}) = \frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}^{*'}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) + \boldsymbol{o}_P(1). \tag{G.119}$$

$\square$

Proof of Lemma F.2(viii). First, expanding,

$$\begin{aligned}
\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) &= \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}(\mathfrak{Z}_1\boldsymbol{\varepsilon}')\\ \vdots \\ \mathrm{tr}(\mathfrak{Z}_P\boldsymbol{\varepsilon}')\end{pmatrix} - \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}(\bar{\mathfrak{Z}}_1\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')\\ \vdots \\ \mathrm{tr}(\bar{\mathfrak{Z}}_P\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')\end{pmatrix} \\
&\quad - \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')\\ \vdots \\ \mathrm{tr}((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_Q)\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')\end{pmatrix} - \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}(\bar{\mathfrak{Z}}_1\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0})\\ \vdots \\ \mathrm{tr}(\bar{\mathfrak{Z}}_P\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0})\end{pmatrix} \\
&\quad - \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0})\\ \vdots \\ \mathrm{tr}((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_Q)\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0})\end{pmatrix} + \frac{1}{\sqrt{nT}}\begin{pmatrix}\mathrm{tr}(\mathfrak{Z}_1\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0})\\ \vdots \\ \mathrm{tr}(\mathfrak{Z}_P\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0})\end{pmatrix} \\
&=: \frac{1}{\sqrt{nT}}(\boldsymbol{l}_1 + ... + \boldsymbol{l}_6).
\end{aligned} \tag{G.120}$$

Consider the 6 terms on the right-hand side of (G.120). First,

$$\mathbb{E}\left[||\boldsymbol{l}_1||_2^2\right] = \mathbb{E}\left[\sum_{p=1}^{P}\mathrm{tr}(\mathfrak{Z}_p\boldsymbol{\varepsilon}')^2\right] = O(PnT) \tag{G.121}$$

using Lemma A.2(v), with $\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}$ replaced by an identity matrix. The implies $||\boldsymbol{l}_1||_2 = O_P(\sqrt{PnT})$. For the next term,

$$\begin{aligned}
||\boldsymbol{l}_2||_2^2 = \sum_{p=1}^{P}\mathrm{tr}(\bar{\mathfrak{Z}}_p\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')^2 &= \frac{1}{T^2}\sum_{p=1}^{P}\mathrm{tr}\left(\boldsymbol{F}^0\left(\frac{1}{T}\boldsymbol{F}^{0'}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\bar{\mathfrak{Z}}_p\right)^2 \\
&\le \frac{1}{T^2}R^2||\boldsymbol{F}^0||_2^2\left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0'}\boldsymbol{F}^0\right)^{-1}\right|\right|_2^2\sum_{p=1}^{P}||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\bar{\mathfrak{Z}}_p||_2^2 \\
&= O_P(PnT)
\end{aligned} \tag{G.122}$$

using Lemma F.1(vii). Therefore $||\boldsymbol{l}_2||_2 = O_P(\sqrt{PnT})$. Analogous steps can be used to show $||\boldsymbol{l}_4||_2 = O_P(\sqrt{PnT})$. For $\boldsymbol{l}_3$,

$$||\boldsymbol{l}_3||_2^2 = \sum_{p=1}^{P}\mathrm{tr}((\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}')^2 \le (R^0)^2||\boldsymbol{\varepsilon}||_2^2||\boldsymbol{P}_{\boldsymbol{F}^0}||_2^2\sum_{p=1}^{P}||\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p||_2^2 = O_P(P\max\{n^2, T^2\})$$

using Lemma F.1(v). Therefore $||\boldsymbol{l}_3||_2 = O_P(\sqrt{P}\max\{n,T\})$. Similar steps can be used to establish that $||\boldsymbol{l}_5||_2 = O_P(\sqrt{P}\max\{n,T\})$. Finally,

$$||\boldsymbol{l}_6||_2^2 = \sum_{p=1}^{P} \mathrm{tr}(\mathfrak{Z}_p \boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0})^2 \leq (R^0)^2||\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2^2 \sum_{p=1}^{P} ||\mathfrak{Z}_p||_2^2 = O_P(PnT) \tag{G.123}$$

as

$$\begin{aligned}
||\boldsymbol{P}_{\boldsymbol{F}^0}\boldsymbol{\varepsilon}'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2 &= \left|\left|\frac{1}{nT}\boldsymbol{F}^0\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0\left(\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\boldsymbol{\Lambda}^{0\prime}\right|\right|_2 \\
&\leq \frac{1}{nT}||\boldsymbol{F}^0||_2||\boldsymbol{\Lambda}^0||_2\left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2 ||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0||_2\left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2 \\
&= O_P(1)
\end{aligned} \tag{G.124}$$

using Lemmas A.2(iii) and F.1(ii). Combining all the above,

$$\frac{1}{\sqrt{nT}}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) = \boldsymbol{O}_P(\sqrt{P}), \tag{G.125}$$

using Assumption 6.4. □

Proof of Lemma F.2(ix). Consider the (p,p')-th element of the matrix $\mathbb{Z}'\mathbb{Z} - \boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}$,

$$\begin{aligned}
&\mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_p\boldsymbol{M}_{\boldsymbol{F}^0} + (\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p))'(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\bar{\mathfrak{Z}}_{p'}\boldsymbol{M}_{\boldsymbol{F}^0} + (\mathfrak{Z}_{p'} - \bar{\mathfrak{Z}}_{p'})) - \mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathfrak{Z}_{p'}\boldsymbol{M}_{\boldsymbol{F}^0}\mathfrak{Z}_p') \\
&= \mathrm{tr}((\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_{p'} - \bar{\mathfrak{Z}}_{p'})) + \mathrm{tr}(\boldsymbol{P}_{\boldsymbol{F}^0}(\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'(\mathfrak{Z}_{p'} - \bar{\mathfrak{Z}}_{p'})) \\
&\quad + \mathrm{tr}(\boldsymbol{P}_{\boldsymbol{F}^0}(\mathfrak{Z}_p - \bar{\mathfrak{Z}}_p)'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_{p'} - \bar{\mathfrak{Z}}_{p'})).
\end{aligned} \tag{G.126}$$

Thus,

$$\begin{aligned}
&\mathbb{Z}'\mathbb{Z} - \boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} \\
= &\ \big(\mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)), ..., \mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P))\big)'\big(\mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)), ..., \mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P))\big) \\
&+ \big(\mathrm{vec}((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)\boldsymbol{P}_{\boldsymbol{F}^0}), ..., \mathrm{vec}((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)\boldsymbol{P}_{\boldsymbol{F}^0})\big)'\big(\mathrm{vec}((\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)\boldsymbol{P}_{\boldsymbol{F}^0}), ..., \mathrm{vec}((\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)\boldsymbol{P}_{\boldsymbol{F}^0})\big) \\
&- \big(\mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)\boldsymbol{P}_{\boldsymbol{F}^0}), ..., \mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)\boldsymbol{P}_{\boldsymbol{F}^0})\big)' \\
&\times \big(\mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_1 - \bar{\mathfrak{Z}}_1)\boldsymbol{P}_{\boldsymbol{F}^0}), ..., \mathrm{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_P - \bar{\mathfrak{Z}}_P)\boldsymbol{P}_{\boldsymbol{F}^0})\big) \\
=: &\ \boldsymbol{L}_1 + \boldsymbol{L}_2 + \boldsymbol{L}_3,
\end{aligned} \tag{G.127}$$

and hence

$$
\begin{aligned}
&\frac{1}{nT}\mathbb{E}\left[\left|\left|\mathbb{Z}'\mathbb{Z} - \mathcal{Z}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathcal{Z}\right|\right|_2\right] \\
&\leq \frac{1}{nT}\mathbb{E}\left[||\boldsymbol{L}_1||_2\right] + \frac{1}{nT}\mathbb{E}\left[||\boldsymbol{L}_2||_2\right] + \frac{1}{nT}\mathbb{E}\left[||\boldsymbol{L}_3||_2\right]. \qquad \text{(G.128)}
\end{aligned}
$$

Consider the first term in (G.128)

$$
\begin{aligned}
\frac{1}{nT}\mathbb{E}\left[||\boldsymbol{L}_1||_2\right] &= \frac{1}{nT}\mathbb{E}\left[||\left(\operatorname{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_1-\bar{\mathfrak{Z}}_1)),...,\operatorname{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_P-\bar{\mathfrak{Z}}_P))\right)||_2^2\right] \\
&\leq \frac{1}{nT}\mathbb{E}\left[||\left(\operatorname{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_1-\bar{\mathfrak{Z}}_1)),...,\operatorname{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_P-\bar{\mathfrak{Z}}_P))\right)||_F^2\right] \\
&= \frac{1}{nT}\sum_{p=1}^{P}\mathbb{E}\left[||\operatorname{vec}(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_p-\bar{\mathfrak{Z}}_p))||_2^2\right] \\
&= \frac{1}{nT}\sum_{p=1}^{P}\mathbb{E}\left[||(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}(\mathfrak{Z}_p-\bar{\mathfrak{Z}}_p))||_F^2\right] \\
&\leq \frac{\sqrt{R}}{nT}\sum_{p=1}^{P}\mathbb{E}\left[||\mathfrak{Z}_p-\bar{\mathfrak{Z}}_p||_2^2\right] \\
&= O\left(\frac{P}{\min\{n,T\}}\right) = o(1), \qquad \text{(G.129)}
\end{aligned}
$$

using Lemma F.1(v). Similarly for $||\boldsymbol{L}_2||_2$ and $||\boldsymbol{L}_3||_2$ which gives the result. $\square$

Proof of Lemma F.3(i).

$$
\begin{aligned}
\boldsymbol{B}_1 &= \begin{pmatrix} \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})) - \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_1\boldsymbol{G}_1) & \cdots & \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})) - \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_1\boldsymbol{G}_Q) \\ \vdots & \ddots & \vdots \\ \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_1(\bar{\boldsymbol{\rho}})) - \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_Q\boldsymbol{G}_1) & \cdots & \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_Q(\bar{\boldsymbol{\rho}})) - \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_Q\boldsymbol{G}_Q) \end{pmatrix} + \boldsymbol{B}_1^* \\
&=: \boldsymbol{B}^{**} + \boldsymbol{B}_1^*.
\end{aligned}
$$

First note that, by adding and subtracting,

$$
\begin{aligned}
\frac{1}{n}\operatorname{tr}(\boldsymbol{G}_q(\bar{\boldsymbol{\rho}})\boldsymbol{G}_{q'}(\bar{\boldsymbol{\rho}})) - \frac{1}{n}\operatorname{tr}(\boldsymbol{G}_q\boldsymbol{G}_{q'}) &= \frac{1}{n}\operatorname{tr}(\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})\boldsymbol{W}_q\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})\boldsymbol{W}_{q'} - \boldsymbol{S}^{-1}\boldsymbol{W}_q\boldsymbol{S}^{-1}\boldsymbol{W}_{q'}) \\
&= \frac{1}{n}\operatorname{tr}(\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})(\boldsymbol{I}_n - \boldsymbol{S}(\bar{\boldsymbol{\rho}})\boldsymbol{S}^{-1})\boldsymbol{W}_q\boldsymbol{S}^{-1}\boldsymbol{W}_{q'}) \\
&\quad - \frac{1}{n}\operatorname{tr}(\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})\boldsymbol{W}_q\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})(\boldsymbol{I}_n - \boldsymbol{S}(\bar{\boldsymbol{\rho}})\boldsymbol{S}^{-1})\boldsymbol{W}_{q'}) \\
&=: \frac{1}{n}\boldsymbol{B}_1^{**} + \frac{1}{n}\boldsymbol{B}_2^{**}. \qquad \text{(G.130)}
\end{aligned}
$$

With $||\boldsymbol{B}^{**}||_2 \leq \frac{1}{n}||\boldsymbol{B}_1^{**}||_F + \frac{1}{n}||\boldsymbol{B}_2^{**}||_F$

$$
\begin{aligned}
||\boldsymbol{B}_1^{**}||_F^2 &= \sum_{q=1}^{Q}\sum_{q'=1}^{Q} \operatorname{tr}(\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})(\boldsymbol{I}_n - \boldsymbol{S}(\bar{\boldsymbol{\rho}})\boldsymbol{S}^{-1})\boldsymbol{W}_q\boldsymbol{S}^{-1}\boldsymbol{W}_{q'})^2 \\
&\leq \sum_{q=1}^{Q}\sum_{q'=1}^{Q} n^2||\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})||_2^2||\boldsymbol{I}_n - \boldsymbol{S}\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})||_2^2||\boldsymbol{W}_q||_2^2||\boldsymbol{S}^{-1}||_2^2||\boldsymbol{W}_{q'}||_2^2 \\
&= O_P(Q^3 n^2||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2^2), \qquad \text{(G.131)}
\end{aligned}
$$

using the inequality $|\operatorname{tr}(\boldsymbol{B})| \leq \operatorname{rank}(\boldsymbol{B})||\boldsymbol{B}||_2$, the fact that an $n \times n$ matrix $\boldsymbol{B}$ can have rank no greater then n, and, by the same steps as those in the proof of Lemma A.2(viii), $||\boldsymbol{I}_n - \boldsymbol{S}\boldsymbol{S}^{-1}(\bar{\boldsymbol{\rho}})|| = O_P(\sqrt{Q}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$ due to the fact that $\bar{\boldsymbol{\rho}} - \boldsymbol{\rho}^0 = w\hat{\boldsymbol{\rho}} + (1-w)\boldsymbol{\rho}^0 - \boldsymbol{\rho}^0 = w(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0)$ whereby $||\bar{\boldsymbol{\rho}} - \boldsymbol{\rho}^0||_2 \leq |w|||\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0||_2 \leq ||\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0||_2$, with $0 < w < 1$. Thus $\frac{1}{n}||\boldsymbol{B}_1^{**}||_F = O_P(Q^{1.5}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$. Similar steps show $\frac{1}{n}||\boldsymbol{B}_2^{**}||_F = O_P(Q^{1.5}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$. This gives the final result $\boldsymbol{B}_1 = \boldsymbol{B}_1^* + O_P(Q^{1.5}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)$. □

Proof of Lemma F.3(ii). By adding and subtracting, $\boldsymbol{B}_2 = \boldsymbol{B}_2^* + (\boldsymbol{B}_2 - \boldsymbol{B}_2^*)$. Let $\boldsymbol{B}_2^{**} := nT\,(\boldsymbol{B}_2 - \boldsymbol{B}_2^*)$. One has

$$
\begin{aligned}
\mathbb{E}\left[||\boldsymbol{B}_2^{**}||_F^2\right] &= \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{q'=1}^{Q}(\operatorname{tr}((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{G}_{q'}\boldsymbol{\varepsilon}) - T\sigma_0^2\operatorname{tr}(\boldsymbol{G}_q'\boldsymbol{G}_{q'}))^2\right] \\
&= \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{q'=1}^{Q}(\operatorname{tr}((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{G}_{q'}\boldsymbol{\varepsilon}))^2\right] - 2\sum_{q=1}^{Q}\sum_{q'=1}^{Q} T\sigma_0^2\operatorname{tr}(\boldsymbol{G}_q'\boldsymbol{G}_{q'})\mathbb{E}\left[\operatorname{tr}((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{G}_{q'}\boldsymbol{\varepsilon})\right] \\
&\quad + \sum_{q=1}^{Q}\sum_{q'=1}^{Q}(T\sigma_0^2\operatorname{tr}(\boldsymbol{G}_q'\boldsymbol{G}_{q'}))^2. \qquad \text{(G.132)}
\end{aligned}
$$

First, $\mathbb{E}\left[\operatorname{tr}((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{G}_{q'}\boldsymbol{\varepsilon})\right] = T\sigma_0^2\operatorname{tr}(\boldsymbol{G}_q'\boldsymbol{G}_{q'})$. Second,

$$
\begin{aligned}
\mathbb{E}\left[(\operatorname{tr}((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{G}_{q'}\boldsymbol{\varepsilon}))^2\right] &= (\mathcal{M}_\varepsilon^4 - 3\sigma_0^4)\sum_{i=1}^{nT}(\boldsymbol{I}_T \otimes \boldsymbol{G}_q'\boldsymbol{G}_{q'})_{ii}^2 \\
&\quad + \sigma_0^4\left(\operatorname{tr}(\boldsymbol{I}_T \otimes \boldsymbol{G}_q'\boldsymbol{G}_{q'})\right)^2 + 2\operatorname{tr}((\boldsymbol{I}_T \otimes \boldsymbol{G}_q'\boldsymbol{G}_{q'})(\boldsymbol{I}_T \otimes \boldsymbol{G}_q'\boldsymbol{G}_{q'})'). \qquad .
\end{aligned}
$$

Therefore (G.132) becomes

$$
\mathbb{E}\left[||\boldsymbol{B}_2^{**}||_F^2\right] = (\mathcal{M}_\varepsilon^2 - 3\sigma_0^4)\sum_{i=1}^{nT}(\boldsymbol{I}_T \otimes \boldsymbol{G}_q'\boldsymbol{G}_{q'})_{ii}^2 + 2\sigma_0^4\operatorname{tr}((\boldsymbol{I}_T \otimes \boldsymbol{G}_q'\boldsymbol{G}_{q'})(\boldsymbol{I}_T \otimes \boldsymbol{G}_q'\boldsymbol{G}_{q'})').
$$

Observe that $\text{tr}((\boldsymbol{I}_T \otimes \boldsymbol{G}_q' \boldsymbol{G}_{q'})(\boldsymbol{I}_T \otimes \boldsymbol{G}_q' \boldsymbol{G}_{q'})') = \text{tr}(\boldsymbol{I}_T \otimes (\boldsymbol{G}_q' \boldsymbol{G}_{q'} \boldsymbol{G}_{q'}' \boldsymbol{G}_q)) = T \text{tr}(\boldsymbol{G}_q' \boldsymbol{G}_{q'} \boldsymbol{G}_{q'}' \boldsymbol{G}_q) = T ||\boldsymbol{G}_q' \boldsymbol{G}_{q'}||_F^2 \leq nT ||\boldsymbol{G}_q' \boldsymbol{G}_{q'}||_2^2$. Similarly $\sum_{i=1}^{nT} (\boldsymbol{I}_T \otimes \boldsymbol{G}_q' \boldsymbol{G}_{q'})_{ii}^2 \leq nT ||\boldsymbol{G}_q' \boldsymbol{G}_{q'}||_2^2$. Combining all the above results yields $\mathbb{E}\left[||\boldsymbol{B}_2^{**}||_F^2\right] = O(Q^2 nT)$, from which the result follows. $\square$

Proof of Lemma F.3(iii). For brevity, recall the definition $\boldsymbol{\mathcal{B}}_3 := \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{Z}_t^*)' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{Z}_t^*$ from the proof of Lemma A.4. It is straightforward to show that $\boldsymbol{\mathcal{B}}_3 - \boldsymbol{\mathfrak{H}}$ is equivalent to

$$\boldsymbol{\mathcal{B}}_3 - \boldsymbol{\mathfrak{H}} = \frac{1}{nT} \begin{pmatrix} \text{tr}((\boldsymbol{W}_1 \boldsymbol{Y})' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{W}_1 \boldsymbol{Y} \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}((\boldsymbol{W}_1 \boldsymbol{Y})' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{\mathcal{X}}_K \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\boldsymbol{\mathcal{X}}_K)' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{W}_1 \boldsymbol{Y} \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}(\boldsymbol{\mathcal{X}}_K' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{\mathcal{X}}_K \boldsymbol{M}_{\boldsymbol{F}^0}) \end{pmatrix}. \tag{G.133}$$

By substituting in the true DGP, using the fact that $\boldsymbol{M}_{\boldsymbol{F}^0} \boldsymbol{F}^0 = \boldsymbol{0}_{T \times R^0}$, and adding and subtracting terms,

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_3 - \boldsymbol{\mathfrak{H}} &= \frac{1}{nT} \begin{pmatrix} \text{tr}(\mathfrak{Z}_1'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_1 \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}(\mathfrak{Z}_1'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_P \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}(\mathfrak{Z}_P'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_1 \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}(\mathfrak{Z}_P'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_P \boldsymbol{M}_{\boldsymbol{F}^0}) \end{pmatrix} \\
&+ \frac{1}{nT} \left(\begin{pmatrix} \text{tr}((\boldsymbol{G}_1 \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_1 \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}((\boldsymbol{G}_Q \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_1 \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\boldsymbol{G}_1 \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_P \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}((\boldsymbol{G}_Q \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_P \boldsymbol{M}_{\boldsymbol{F}^0}) \end{pmatrix} \quad \boldsymbol{0}_{P \times K} \right) \\
&+ \frac{1}{nT} \left(\begin{pmatrix} \text{tr}((\boldsymbol{G}_1 \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_1 \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}((\boldsymbol{G}_Q \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_1 \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\boldsymbol{G}_1 \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_P \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}((\boldsymbol{G}_Q \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_P \boldsymbol{M}_{\boldsymbol{F}^0}) \end{pmatrix} \quad \boldsymbol{0}_{P \times K} \right)' \\
&+ \frac{1}{nT} \begin{pmatrix} \begin{pmatrix} \text{tr}((\boldsymbol{G}_1 \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \boldsymbol{G}_1 \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}((\boldsymbol{G}_1 \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \boldsymbol{G}_Q \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\boldsymbol{G}_Q \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \boldsymbol{G}_1 \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}(\boldsymbol{G}_Q \boldsymbol{\varepsilon}(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \boldsymbol{G}_Q \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) \end{pmatrix} & \boldsymbol{0}_{Q \times K} \\ \boldsymbol{0}_{K \times Q} & \boldsymbol{0}_{K \times K} \end{pmatrix} \\
&+ \frac{1}{nT} \begin{pmatrix} \text{tr}((\boldsymbol{W}_1 \boldsymbol{Y})' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{W}_1 \boldsymbol{Y} \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}((\boldsymbol{W}_1 \boldsymbol{Y})' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\mathcal{X}}_K \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\boldsymbol{\mathcal{X}}_K)' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{W}_1 \boldsymbol{Y} \boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \text{tr}(\boldsymbol{\mathcal{X}}_K' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\mathcal{X}}_K \boldsymbol{M}_{\boldsymbol{F}^0}). \end{pmatrix} \\
&=: \boldsymbol{L}_1 + \boldsymbol{L}_2 + \boldsymbol{L}_3 + \boldsymbol{L}_4 + \boldsymbol{L}_5.
\end{aligned} \tag{G.134}$$

For the first term,

$$||\boldsymbol{L}_1||_F^2 = \frac{1}{n^2 T^2} \sum_{p=1}^P \sum_{p'=1}^P \text{tr}(\mathfrak{Z}_p'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0}) \mathfrak{Z}_{p'} \boldsymbol{M}_{\boldsymbol{F}^0})^2$$

$$\leq \frac{1}{n^2T^2}\sum_{p=1}^{P}\sum_{p'=1}^{P}||\mathcal{Z}_p'(\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}-\boldsymbol{P}_{\boldsymbol{\Lambda}^0})||_F^2||\mathcal{Z}_{p'}\boldsymbol{M}_{\boldsymbol{F}^0}||_F^2$$

$$\leq \frac{1}{n^2T^2}\left(\sum_{p=1}^{P}||\mathcal{Z}_p||_F^2\right)\left(\sum_{p'=1}^{P}||\mathcal{Z}_{p'}||_F^2\right)||\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}-\boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2^2||\boldsymbol{M}_{\boldsymbol{F}^0}||_2^2. \tag{G.135}$$

Hence $||\boldsymbol{L}_1||_F = O_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+O_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)$. The same steps can be followed to establish that $||\boldsymbol{L}_2||_F = ||\boldsymbol{L}_3||_F = ||\boldsymbol{L}_4||_F = O_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+O_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)$. For the last term, $\boldsymbol{L}_5$, this can be expanded to yield

$$\begin{aligned}
\boldsymbol{L}_5 &= \frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}} \\
&\quad+\frac{1}{nT}\begin{pmatrix}\begin{pmatrix}\operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_1\boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_1\boldsymbol{M}_{\boldsymbol{F}^0})\\ \vdots & \ddots & \vdots \\ \operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_P\boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_P\boldsymbol{M}_{\boldsymbol{F}^0})\end{pmatrix} & \mathbf{0}_{P\times K}\end{pmatrix} \\
&\quad+\frac{1}{nT}\begin{pmatrix}\begin{pmatrix}\operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_1\boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_1\boldsymbol{M}_{\boldsymbol{F}^0})\\ \vdots & \ddots & \vdots \\ \operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_P\boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\mathcal{Z}_P\boldsymbol{M}_{\boldsymbol{F}^0})\end{pmatrix} & \mathbf{0}_{P\times K}\end{pmatrix}' \\
&\quad+\frac{1}{nT}\begin{pmatrix}\begin{pmatrix}\operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_1\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_Q\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\\ \vdots & \ddots & \vdots \\ \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_1\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) & \cdots & \operatorname{tr}(\boldsymbol{G}_Q\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_Q\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})\end{pmatrix} & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix} \\
&=: \frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}+\boldsymbol{L}_{5.1}+\boldsymbol{L}_{5.2}+\boldsymbol{L}_{5.3}.
\end{aligned} \tag{G.136}$$

Using the independence of the errors from the factors, the loadings and the covariates, it is straightforward to establish that $||\boldsymbol{L}_{5.1}||_F = ||\boldsymbol{L}_{5.2}||_F = O_P(\sqrt{QP}/\sqrt{nT})$. For $\boldsymbol{L}_{5.3}$,

$$\begin{aligned}
\boldsymbol{L}_{5.3} &= \frac{1}{nT}\begin{pmatrix}\begin{pmatrix}\operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{G}_1\boldsymbol{\varepsilon}) & \cdots & \operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{G}_Q\boldsymbol{\varepsilon})\\ \vdots & \ddots & \vdots \\ \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{G}_1\boldsymbol{\varepsilon}) & \cdots & \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{G}_Q\boldsymbol{\varepsilon})\end{pmatrix} & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix} \\
&\quad-\frac{1}{nT}\begin{pmatrix}\begin{pmatrix}\operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_1\boldsymbol{\varepsilon}) & \cdots & \operatorname{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_Q\boldsymbol{\varepsilon})\\ \vdots & \ddots & \vdots \\ \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_1\boldsymbol{\varepsilon}) & \cdots & \operatorname{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_Q\boldsymbol{\varepsilon})\end{pmatrix} & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix}
\end{aligned}$$

$$
-\frac{1}{nT}\begin{pmatrix}\begin{pmatrix}\mathrm{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{G}_1\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) & \cdots & \mathrm{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{G}_Q\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\\ \vdots & \ddots & \vdots\\ \mathrm{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{G}_1\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) & \cdots & \mathrm{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{G}_Q\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\end{pmatrix} & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix}
$$

$$
+\frac{1}{nT}\begin{pmatrix}\begin{pmatrix}\mathrm{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_1\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) & \cdots & \mathrm{tr}((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_Q\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\\ \vdots & \ddots & \vdots\\ \mathrm{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_1\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) & \cdots & \mathrm{tr}((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_Q\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0})\end{pmatrix} & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix}
$$

$$
=: \frac{1}{nT}(\boldsymbol{L}_{5.3.1}+\boldsymbol{L}_{5.3.2}+\boldsymbol{L}_{5.3.3}+\boldsymbol{L}_{5.3.4}). \tag{G.137}
$$

For $\boldsymbol{L}_{5.3.2}$,

$$
\begin{aligned}
||\boldsymbol{L}_{5.3.2}||_F^2 &= \sum_{q=1}^{Q}\sum_{q'=1}^{Q}\mathrm{tr}((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_{q'}\boldsymbol{\varepsilon})^2\\
&\le \sum_{q=1}^{Q}\sum_{q'=1}^{Q}(R^0)^2||(\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{G}_{q'}\boldsymbol{\varepsilon}||_2^2\\
&\le \sum_{q=1}^{Q}\sum_{q'=1}^{Q}R^2||\boldsymbol{G}_q||_2^2||\boldsymbol{G}_{q'}||_2^2||\boldsymbol{\varepsilon}||_2^4||\boldsymbol{P}_{\boldsymbol{\Lambda}^0}||_2^2\\
&= (R^0)^2||\boldsymbol{\varepsilon}||_2^4\sum_{q=1}^{Q}||\boldsymbol{G}_q||_2^2\sum_{q'=1}^{Q}||\boldsymbol{G}_{q'}||_2^2\\
&= O_P(Q^2(\max\{n,T\})^2),
\end{aligned} \tag{G.138}
$$

which implies $\frac{1}{nT}||\boldsymbol{L}_{5.3.2}||_F = O_P\left(\frac{Q}{\min\{n,T\}}\right)$. The same steps can be followed to establish that $\frac{1}{nT}||\boldsymbol{L}_{5.3.3}||_F$ and $\frac{1}{nT}||\boldsymbol{L}_{5.3.4}||_F$ have the same probability order. Combining these results and using Lemma F.3(ii) gives the final result

$$
\begin{aligned}
\boldsymbol{\mathcal{B}}_3-\mathcal{H} &= \frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}+\sigma_0^2\begin{pmatrix}\begin{pmatrix}\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1'\boldsymbol{G}_1) & \cdots & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1'\boldsymbol{G}_Q)\\ \vdots & \ddots & \vdots\\ \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q'\boldsymbol{G}_1) & \cdots & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q'\boldsymbol{G}_Q)\end{pmatrix} & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix}\\
&\quad+\boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\right)+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right).
\end{aligned} \tag{G.139}
$$

□

Proof of Lemma F.3(iv). Recalling $\boldsymbol{Z}_t^* := (\boldsymbol{W}_1\boldsymbol{y}_t, ..., \boldsymbol{W}_Q\boldsymbol{y}_t, \boldsymbol{X}_t)$, the $P \times 1$ vector $\boldsymbol{\mathcal{B}}_4 := \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)$ can be expanded as

$$\begin{aligned}\boldsymbol{\mathcal{B}}_4 &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t + \begin{pmatrix}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_1(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t))'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)\\ \vdots \\ \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_Q(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t))'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)\\ \boldsymbol{0}_{K\times 1}\end{pmatrix}\\ &=: \boldsymbol{\mathcal{B}}_{4.1} + \boldsymbol{\mathcal{B}}_{4.2} + \boldsymbol{\mathcal{B}}_{4.3}. \end{aligned} \tag{G.140}$$

Note that $\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0 = \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0 - \hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})$, where $\boldsymbol{H}^* := \frac{1}{nT}\boldsymbol{F}^{0'}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Pi}^{-1}$ is a $R^0 \times R$ matrix ($R^0 = R$ by Assumption 6.5) and $\boldsymbol{\Pi}$ is a diagonal $R \times R$ matrix containing the largest R eigenvalues of $\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{e}_t\boldsymbol{e}_t'$ along its diagonal. Therefore, $\boldsymbol{\mathcal{B}}_{4.1} = \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0 - \hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})\boldsymbol{f}_t^0$, which, by decomposition (G.17) in the proof of Lemma F.2(i), can be expanded as

$$\begin{aligned}\boldsymbol{\mathcal{B}}_{4.1} &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-(\boldsymbol{P}_1 + ... + \boldsymbol{P}_6 + \boldsymbol{P}_8 + ... + \boldsymbol{P}_{25})\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0\\ &=: \boldsymbol{\mathcal{B}}_{4.1.1} + ... + \boldsymbol{\mathcal{B}}_{4.1.6} + \boldsymbol{\mathcal{B}}_{4.1.8} + ... + \boldsymbol{\mathcal{B}}_{4.1.25},\end{aligned} \tag{G.141}$$

where $\boldsymbol{\Sigma}^* := \left(\frac{1}{nT}\boldsymbol{F}^{0'}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\hat{\boldsymbol{\Lambda}}\right)^{-1}$ is well defined by Assumption 6.3 and Lemma F.2(ii), with $||\boldsymbol{\Sigma}^*||_2 = O_P(1)$. The probability order of the 24 terms in (G.141) is now examined, though for brevity derivations for similar terms are omitted. Starting with the first term,

$$\boldsymbol{\mathcal{B}}_{4.1.1} = \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0)'\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0),$$

since $(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\boldsymbol{Z}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$ is a scalar. Now

$$\begin{aligned}&\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0)'\boldsymbol{Z}_\tau\right|\right|_2\\ &\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\sum_{t=1}^{T}\sum_{\tau=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2||\boldsymbol{f}_t^0||_2||\boldsymbol{Z}_\tau||_2\\ &\leq \frac{1}{nT}\frac{1}{\sqrt{nT}}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}\left(\frac{1}{nT}\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})||_2^2\right)^{\frac{1}{2}}\\ &\quad\times\left(\sum_{t=1}^{T}||\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}}\left(\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau||_2^2\right)^{\frac{1}{2}} = O_P(P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)\end{aligned} \tag{G.142}$$

using Lemmas A.2(iii), A.2(iv), A.3(i) and F.1(i). Thus, $\boldsymbol{\mathcal{B}}_{4.1.1} = \boldsymbol{O}_P(P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$. For term $\boldsymbol{\mathcal{B}}_{4.1.2}$,

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.2} &= \frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{Z}_t' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \left(-\frac{1}{nT}\sum_{\tau=1}^{T} \boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\boldsymbol{f}_\tau^{0\prime}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^* \right) \boldsymbol{f}_t^0 \\
&= -\frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{Z}_t' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \left(\frac{1}{T}\sum_{\tau=1}^{T} \boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\boldsymbol{f}_\tau^{0\prime} \left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\right) \left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\right)^{-1} \left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1} \right) \boldsymbol{f}_t^0 \\
&= -\frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{Z}_t' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \left(\frac{1}{T}\sum_{\tau=1}^{T} \boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\boldsymbol{f}_\tau^{0\prime} \left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1} \boldsymbol{f}_t^0 \right).
\end{aligned}$$

Letting $\varpi_{\tau t}^0 := \boldsymbol{f}_\tau^{0\prime}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{f}_t^0$,

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.2} &= -\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T} \varpi_{\tau t}^0 \boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) \\
&= \left(\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T} \varpi_{\tau t}^0 \boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0).
\end{aligned}$$

Next

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.3} &= \frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{Z}_t' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \left(-\frac{1}{nT}\sum_{\tau=1}^{T} \boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^* \right) \boldsymbol{f}_t^0 \\
&= \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T} \boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \qquad \text{(G.143)}
\end{aligned}$$

since $\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$ is a scalar. Now,

$$\begin{aligned}
&\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T} \boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right|\right|_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{\Sigma}^*||_2||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2\sum_{t=1}^{T}\sum_{\tau=1}^{T}||\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{f}_t^0||_2||\boldsymbol{Z}_t||_2||\boldsymbol{Z}_\tau||_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{\Sigma}^*||_2||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2\left(\sum_{\tau=1}^{T}||\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}\left(\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau||_2^2\right)^{\frac{1}{2}} \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{\Sigma}^*||_2||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\varepsilon}||_2||\hat{\boldsymbol{\Lambda}}||_F||\boldsymbol{F}^0||_F\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}\left(\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right) \qquad \text{(G.144)}
\end{aligned}$$

using Lemmas A.2(iii), A.2(iv), A.2(vi) and F.1(i). Thus, $\boldsymbol{\mathcal{B}}_{4.1.3} = \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$. For $\boldsymbol{\mathcal{B}}_{4.1.4}$,

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.4} &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0 \\
&= \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \quad \text{(G.145)}
\end{aligned}$$

since $\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$ is a scalar. Now,

$$\begin{aligned}
&\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\right|\right|_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\sum_{t=1}^{T}\sum_{\tau=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{Z}_\tau||_2||\boldsymbol{f}_\tau^0||_2||\boldsymbol{f}_t^0||_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2\left|\left|\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\right|\right|_2||\boldsymbol{\Lambda}^0||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\left(\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{F}^0||_F^2 \\
&= O_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) \quad \text{(G.146)}
\end{aligned}$$

using Lemmas A.2(i), A.2(iii), A.2(iv), A.2(viii) and F.1(i). Thus, $\boldsymbol{\mathcal{B}}_{4.1.4} = \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$. For $\boldsymbol{\mathcal{B}}_{4.1.5}$,

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.5} &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0 \\
&= \left(\frac{1}{nT}\frac{1}{nT}\sum_{\tau=1}^{T}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0), \quad \text{(G.147)}
\end{aligned}$$

since $\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$ is a scalar. Now,

$$\begin{aligned}
&\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{\tau=1}^{T}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\left(\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\right|\right|_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\sum_{q=1}^{Q}(\rho_q^0 - \hat{\rho}_q)\boldsymbol{G}_q||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\sum_{\tau=1}^{T}\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{Z}_\tau||_2||\boldsymbol{\varepsilon}_\tau||_2||\boldsymbol{f}_t^0||_2
\end{aligned}$$

$$\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\left(\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{\varepsilon}||_F||\boldsymbol{F}^0||_F$$
$$= O_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2) \tag{G.148}$$

using Lemmas A.2(i), A.2(iii), A.2(iv), A.2(vi), A.2(viii) and F.1(i). Thus, $\boldsymbol{\mathcal{B}}_{4.1.5} = \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)$. By similar steps it is straightforward to establish that $\boldsymbol{\mathcal{B}}_{4.1.6} = \boldsymbol{O}_P(\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)$. Next,

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.8} &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0\\
&= -\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})\boldsymbol{f}_\tau^0\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0.
\end{aligned} \tag{G.149}$$

Now,

$$\begin{aligned}
&\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})\boldsymbol{f}_\tau^0\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\right|\right|_2\\
&\leq \frac{1}{nT}\frac{1}{nT}\frac{1}{\hat{\sigma}^2}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{\Sigma}^*||_2\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2\left|\left|\sum_{\tau=1}^{T}\boldsymbol{f}_\tau^0\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\right|\right|_2||\boldsymbol{f}_t^0||_2\\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{\Sigma}^*||_2\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{F}^0||_F||\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}||_2\\
&= O_P\left(\frac{\sqrt{P}}{\min\{n\sqrt{T},T^{1.5}\}}\right)+O_P\left(\frac{P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n\sqrt{T},T^{1.5}\}}\right)\\
&\quad + O_P\left(\frac{\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2}{\min\{\sqrt{nT},T\}}\right)+O_P\left(\frac{Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^3}{\sqrt{T}}\right)
\end{aligned} \tag{G.150}$$

using Lemmas A.2(iii), A.2(iv), F.2(i) and F.2(iii).

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.9} &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0\\
&= \frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{F}^{0'}\boldsymbol{F}^0\left(\sum_{q=1}^{Q}(\hat{\rho}_q-\rho_q^0)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\\
&= \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})\boldsymbol{F}^{0'}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}(\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n)'\left(\mathrm{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\mathrm{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\right)\right)
\end{aligned}$$

$$\times(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0). \tag{G.151}$$

Note that

$$\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})\boldsymbol{F}^{0'}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}(\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n)'\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\right|\right|_2$$
$$\leq\frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{F}^0||_2^2\big|\big|\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\big|\big|_2||\boldsymbol{\Lambda}^0||_2\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2$$
$$\times||\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n||_2$$
$$\leq\frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{F}^0||_2^2\big|\big|\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\big|\big|_2||\boldsymbol{\Lambda}^0||_2||\boldsymbol{\Sigma}^*||_2$$
$$\times\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{f}_t^0||_2$$
$$\leq\frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{F}^0||_F^3\big|\big|\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\big|\big|_2||\boldsymbol{\Lambda}^0||_2||\boldsymbol{\Sigma}^*||_2$$
$$\times\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}=O_P(Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+O_P\left(\frac{\sqrt{QP}}{\min\{n,T\}}\right), \tag{G.152}$$

where the second inequality uses the fact that $||\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n||_2=||\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}||_2$ and the last line follows by Lemmas A.2(iii), A.2(iv), F.2(i) and F.1(iii). Hence the result $\boldsymbol{\mathcal{B}}_{4.1.9}=\left(\boldsymbol{O}_P(Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+\boldsymbol{O}_P\left(\frac{\sqrt{QP}}{\min\{n,T\}}\right)\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$. For term $\boldsymbol{\mathcal{B}}_{4.1.10}$,

$$\boldsymbol{\mathcal{B}}_{4.1.10}=\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0$$
$$=\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'\left(\sum_{q=1}^{Q}(\hat{\rho}_q-\rho_q^0)\boldsymbol{G}_q\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$$
$$=\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'(\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n)'\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)$$
$$\times(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0). \tag{G.153}$$

Now,

$$\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1})\boldsymbol{F}^{0'}\boldsymbol{\varepsilon}'(\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n)'\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\right|\right|_2$$
$$\leq\frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\Lambda}^0-\hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2||\boldsymbol{\Sigma}^*||_2\big|\big|\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\big|\big|_2$$

$$\times \sum_{t=1}^{T} ||\boldsymbol{Z}_t||_2 ||\boldsymbol{f}_t^{0\prime}||_2$$

$$\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\Lambda}^0 - \hat{\boldsymbol{\Lambda}}\boldsymbol{H}^{*-1}||_2||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2||\boldsymbol{F}^0||_F||\boldsymbol{\Sigma}^*||_2||\boldsymbol{\varepsilon}||_2\big|\big|\big(\mathrm{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}), ..., \mathrm{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\big|\big|_2$$

$$\times \left(\sum_{t=1}^{T} ||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}$$

$$= O_P\left(\frac{Q\sqrt{P}||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2}{\sqrt{T}}\right) + O_P\left(\frac{\sqrt{QP}}{\min\{\sqrt{nT}, T\}}\right) \tag{G.154}$$

using Lemmas A.2(iii), A.2(iv), F.1(ii), F.1(iii) and F.2(i). This gives the result $\boldsymbol{\mathcal{B}}_{4.1.10} = \left(\boldsymbol{O}_P\left(\frac{Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{T}}\right) + \boldsymbol{O}_P\left(\frac{\sqrt{QP}}{\min\{\sqrt{nT},T\}}\right)\right)(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0)$. Next

$$\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.11} &= \frac{1}{nT}\sum_{t=1}^{T} \boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_\tau(\boldsymbol{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}))'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0 \\
&= \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau(\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0)'\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0).
\end{aligned} \tag{G.155}$$

Now,

$$\begin{aligned}
&\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau(\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0)'\boldsymbol{Z}_\tau\right|\right|_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\sum_{t=1}^{T}\sum_{\tau=1}^{T}||\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau||_2||\boldsymbol{f}_t^0||_2||\boldsymbol{Z}_\tau||_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}||\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau||_2^2\right)^{\frac{1}{2}}\left(\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}} \\
&\leq \frac{1}{nT}\frac{1}{nT}||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}||_2\left(\sum_{p=1}^{P}||\boldsymbol{\mathcal{Z}}_p||_2^2\right)^{\frac{1}{2}}\left(\sum_{\tau=1}^{T}||\boldsymbol{Z}_\tau||_2^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}||\boldsymbol{f}_t^0||_2^2\right)^{\frac{1}{2}} \\
&= O_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)
\end{aligned} \tag{G.156}$$

using Lemmas A.2(iii), A.2(iv) and F.1(i). Hence, $\boldsymbol{\mathcal{B}}_{4.1.11} = \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)$. For term $\boldsymbol{\mathcal{B}}_{4.1.12}$,

$$\boldsymbol{\mathcal{B}}_{4.1.12} = \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_\tau(\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0$$

$$= -\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_\tau(\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$$

$$= -\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\boldsymbol{f}_\tau^{0'}\left(\frac{1}{nT}\boldsymbol{F}^{0'}\boldsymbol{F}^0\right)^{-1}\boldsymbol{f}_t^0$$

$$= -\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\varpi_{\tau t}^0 \tag{G.157}$$

and using Lemma F.2(iv), $||\boldsymbol{\mathcal{B}}_{4.1.13}||_2 = \boldsymbol{O}_P\left(\frac{\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right)+\boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right)$. For $\boldsymbol{\mathcal{B}}_{4.1.14}$,

$$\boldsymbol{\mathcal{B}}_{4.1.14} = \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\boldsymbol{\varepsilon}_\tau\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0$$

$$= -\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}\left(\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\right)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$$

$$= \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}(\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n)'\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\right)(\boldsymbol{\rho}^0-\hat{\boldsymbol{\rho}}). \tag{G.158}$$

Then,

$$\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{F}^0\boldsymbol{\Lambda}^{0'}(\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n)'\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)\right|\right|_2$$

$$\leq \frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2||\boldsymbol{\Lambda}^0||_2||(\boldsymbol{\Sigma}^*\boldsymbol{f}_t^{0'}\otimes\boldsymbol{I}_n)||_2||\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)||_2$$

$$\leq \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2||\boldsymbol{\Lambda}^0||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{F}^0||_F||\big(\text{vec}(\boldsymbol{G}_1'\hat{\boldsymbol{\Lambda}}),...,\text{vec}(\boldsymbol{G}_Q'\hat{\boldsymbol{\Lambda}})\big)||_2$$

$$= O_P\left(\sqrt{\frac{QP}{T}}\right) \tag{G.159}$$

using Lemmas A.2(ii), A.2(iii), F.1(ii) and F.1(iii). Therefore, $\boldsymbol{\mathcal{B}}_{4.1.14} = \boldsymbol{O}_P\left(\sqrt{\frac{QP}{T}}\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$. Similar steps can be used to show $\boldsymbol{\mathcal{B}}_{4.1.15}$ and $\boldsymbol{\mathcal{B}}_{4.1.16}$ are $\boldsymbol{O}_P\left(\frac{\sqrt{QP}}{\min\{n,T\}}\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$ and $\boldsymbol{O}_P\left(\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\right)$ respectively. Next

$$\boldsymbol{\mathcal{B}}_{4.1.17} = \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0(\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0$$

$$
\begin{aligned}
&= \frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\sum_{q=1}^{Q}(\hat{\rho}_q-\rho_q^0)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0(\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&= \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0).
\end{aligned} \tag{G.160}
$$

For term $\boldsymbol{\mathcal{B}}_{4.1.18}$,

$$
\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.1.18} &= \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-\frac{1}{nT}\sum_{\tau=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0 \\
&= -\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0 \\
&= \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\otimes\boldsymbol{I}_n\right)'\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\Lambda}^0),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\Lambda}^0)\right)\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0).
\end{aligned}
$$

Now,

$$
\begin{aligned}
&\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\otimes\boldsymbol{I}_n\right)'\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\Lambda}^0),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\Lambda}^0)\right)\right|\right|_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\left(\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\otimes\boldsymbol{I}_n\right)||_2\left|\left|\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\Lambda}^0),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\Lambda}^0)\right)\right|\right|_2 \\
&\leq \frac{1}{nT}\frac{1}{nT}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\Lambda}^0),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\Lambda}^0)\right)||_2\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{F}^0||_F \\
&= O_P\left(\sqrt{\frac{QP}{T}}\right),
\end{aligned} \tag{G.161}
$$

using Lemmas A.2(iii), F.1(i) and F.1(ii), and where analogous steps to those in the prof of Lemma F.1(iii) can be used to show $||(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\Lambda}^0),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\Lambda}^0))||_2 = O(\sqrt{Qn})$. Thus, $\boldsymbol{\mathcal{B}}_{4.1.18} = \boldsymbol{O}_P\left(\sqrt{\frac{QP}{T}}\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$. Similar steps can be used to establish that $\boldsymbol{\mathcal{B}}_{4.1.19}$, $\boldsymbol{\mathcal{B}}_{4.1.20}$ and $\boldsymbol{\mathcal{B}}_{4.1.21}$ are $\boldsymbol{O}_P\left(Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$, $\boldsymbol{O}_P\left(\frac{Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{T}}\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$ and $\boldsymbol{O}_P\left(\sqrt{QP}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$, respectively. Now consider term $\boldsymbol{\mathcal{B}}_{4.1.22}$.

$$
\boldsymbol{\mathcal{B}}_{4.1.22} = \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\Big(-\frac{1}{nT}\sum_{\tau=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau(\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\Big)\boldsymbol{f}_t^0
$$

$$= -\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau(\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0)'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$$
$$= \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\varepsilon}_\tau,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\varepsilon}_\tau\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0). \tag{G.162}$$

For term $\boldsymbol{\mathcal{B}}_{4.1.23}$,

$$\boldsymbol{\mathcal{B}}_{4.1.23} = \frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\Big(-\frac{1}{nT}\sum_{\tau=1}^{T}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}_\tau\boldsymbol{\varepsilon}_\tau'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\Big)\boldsymbol{f}_t^0$$
$$= -\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\sum_{q=1}^{Q}(\rho_q^0-\hat{\rho}_q)\boldsymbol{G}_q\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0$$
$$= \left(\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\otimes\boldsymbol{I}_n\right)'\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\varepsilon}),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\varepsilon})\right)\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0). \tag{G.163}$$

Now,

$$\left|\left|\frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(\boldsymbol{\varepsilon}'\hat{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}^*\boldsymbol{f}_t^0\otimes\boldsymbol{I}_n\right)'\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\varepsilon}),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\varepsilon})\right)\right|\right|_2$$
$$\leq \frac{1}{nT}\frac{1}{nT}\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\varepsilon}||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{f}_t^0||_2\big|\big|\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\varepsilon}),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\varepsilon})\right)\big|\big|_2$$
$$\leq \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}||\boldsymbol{Z}_t||_2^2\right)^{\frac{1}{2}}||\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{\varepsilon}||_2||\hat{\boldsymbol{\Lambda}}||_2||\boldsymbol{\Sigma}^*||_2||\boldsymbol{F}^0||_F\big|\big|\left(\mathrm{vec}(\boldsymbol{G}_1\boldsymbol{\varepsilon}),...,\mathrm{vec}(\boldsymbol{G}_Q\boldsymbol{\varepsilon})\right)\big|\big|_2$$
$$= O_P\left(\sqrt{\frac{QP}{\min\{n,T\}}}\right), \tag{G.164}$$

using Lemmas A.2(ii), A.2(iii), F.1(i) and F.1(iv). Therefore, $\boldsymbol{\mathcal{B}}_{4.1.23} = \boldsymbol{O}_P\left(\sqrt{\frac{QP}{\min\{n,T\}}}\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$. For the remaining terms, $\boldsymbol{\mathcal{B}}_{4.1.24}$ and $\boldsymbol{\mathcal{B}}_{4.1.25}$ can both be shown to be $\boldsymbol{O}_P\left(\frac{Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\sqrt{T}}\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$ by similar steps to those for $\boldsymbol{\mathcal{B}}_{4.1.14}$. Collecting all these terms gives

$$\boldsymbol{\mathcal{B}}_{4.1} = \left(\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0) - \frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\varpi_{\tau t}^0$$
$$+ \frac{1}{nT}\frac{1}{nT}\left(\textstyle\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0,\cdots,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)$$

$$
\begin{aligned}
&+\frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\varepsilon}_\tau,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\varepsilon}_\tau\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)\\
&+\boldsymbol{\mathcal{B}}_{4.1.8}+\boldsymbol{\mathcal{B}}_{4.1.13}+\boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2+\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\\
=&\left(\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)-\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\varpi_{\tau t}^{0}\\
&+\frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0,\cdots,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)\\
&+\frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\varepsilon}_\tau,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\varepsilon}_\tau\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)\\
&+\boldsymbol{\Delta}_a+\boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2+\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0),
\end{aligned}
\tag{G.165}
$$

where $\boldsymbol{\Delta}_a \coloneqq \boldsymbol{\mathcal{B}}_{4.1.8}+\boldsymbol{\mathcal{B}}_{4.1.13}$ is $\boldsymbol{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right)+\boldsymbol{O}_P\left(\frac{P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right)+\boldsymbol{O}_P\left(\frac{\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2}{\min\{\sqrt{nT},T\}}\right)+\boldsymbol{O}_P\left(\frac{Q\sqrt{P}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^3}{\sqrt{T}}\right)$. Next,

$$
\begin{aligned}
\boldsymbol{\mathcal{B}}_{4.3}=&\begin{pmatrix}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\\ \vdots\\ \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\\ \boldsymbol{0}_{K\times1}\end{pmatrix}+\begin{pmatrix}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t\\ \vdots\\ \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t\\ \boldsymbol{0}_{K\times1}\end{pmatrix}\\
&+\begin{pmatrix}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_1\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\\ \vdots\\ \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_Q\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0\\ \boldsymbol{0}_{K\times1}\end{pmatrix}+\begin{pmatrix}\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_1\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t\\ \vdots\\ \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{G}_Q\boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t\\ \boldsymbol{0}_{K\times1}\end{pmatrix}\\
&\eqqcolon \boldsymbol{\mathcal{B}}_{4.3.1}+\boldsymbol{\mathcal{B}}_{4.3.2}+\boldsymbol{\mathcal{B}}_{4.3.3}+\boldsymbol{\mathcal{B}}_{4.3.4}.
\end{aligned}
\tag{G.166}
$$

Terms $\boldsymbol{\mathcal{B}}_{4.3.1}$ and $\boldsymbol{\mathcal{B}}_{4.3.3}$ can be written more compactly as $\boldsymbol{\mathcal{B}}_{4.3.1}=\frac{1}{nT}\sum_{t=1}^{T}(\mathfrak{M}_t^1)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0$ and $\boldsymbol{\mathcal{B}}_{4.3.3}=\frac{1}{nT}\sum_{t=1}^{T}(\mathfrak{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0$ where $\mathfrak{M}_t^1 \coloneqq \left(\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0,...,\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0,\boldsymbol{0}_{n\times K}\right)$ and $\mathfrak{M}_t^2 \coloneqq (\boldsymbol{G}_1\boldsymbol{\varepsilon}_t,...,\boldsymbol{G}_Q\boldsymbol{\varepsilon}_t,\boldsymbol{0}_{n\times K})$. Now note that

$$
\begin{aligned}
\left(\sum_{t=1}^{T}||\mathfrak{M}_t^1|_2^2\right)^{\frac{1}{2}}\le\left(\sum_{t=1}^{T}||\mathfrak{M}_t^1|_F^2\right)^{\frac{1}{2}}&=\left(\sum_{q=1}^{Q}||\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F^2\right)^{\frac{1}{2}}\\
&\le||\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}||_F^2\sqrt{Q}\sqrt{\max_{1\le q\le Q}||\boldsymbol{G}_q||_2}=O_P(\sqrt{QnT}),
\end{aligned}
\tag{G.167}
$$

and

$$\left(\sum_{t=1}^{T}||\mathcal{M}_t^2||_2^2\right)^{\frac{1}{2}}=\left(\sum_{t=1}^{T}||\mathcal{M}_t^2||_F^2\right)^{\frac{1}{2}}=\left(\sum_{q=1}^{Q}||\boldsymbol{G}_q\boldsymbol{\varepsilon}||_F^2\right)^{\frac{1}{2}}$$
$$\leq||\boldsymbol{\varepsilon}||_F^2\sqrt{Q}\sqrt{\max_{1\leq q\leq Q}||\boldsymbol{G}_q||_2}=O_P(\sqrt{QnT}),\tag{G.168}$$

since $\left(\sum_{t=1}^{T}||\boldsymbol{B}_t||_2^2\right)^{\frac{1}{2}}\leq\left(\sum_{t=1}^{T}||\boldsymbol{B}_t||_F^2\right)^{\frac{1}{2}}$ for any $n\times m$ matrix $\boldsymbol{B}$ and $\boldsymbol{G}_q$ is UB across q. Using this, terms $\boldsymbol{\mathcal{B}}_{4.3.1}$ and $\boldsymbol{\mathcal{B}}_{4.3.3}$ can be expanded in the same way as $\boldsymbol{\mathcal{B}}_{4.1}$, via the decomposition (G.16) in the proof of Lemma F.2(i), i.e.,

$$\boldsymbol{\mathcal{B}}_{4.3.1}=\frac{1}{nT}\sum_{t=1}^{T}(\mathcal{M}_t^1)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-(\boldsymbol{P}_1+...+\boldsymbol{P}_6+\boldsymbol{P}_8+...+\boldsymbol{P}_{25})\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0$$

and

$$\boldsymbol{\mathcal{B}}_{4.3.3}=\frac{1}{nT}\sum_{t=1}^{T}(\mathcal{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\left(-(\boldsymbol{P}_1+...+\boldsymbol{P}_6+\boldsymbol{P}_8+...+\boldsymbol{P}_{25})\boldsymbol{\Sigma}^*\right)\boldsymbol{f}_t^0.$$

Following analogous steps as those for terms $\boldsymbol{\mathcal{B}}_{4.1.1},...,\boldsymbol{\mathcal{B}}_{4.1.6},\boldsymbol{\mathcal{B}}_{4.1.8},...,\boldsymbol{\mathcal{B}}_{4.25}$ yields the expression (the counterpart to (G.165))

$$\begin{aligned}
&\boldsymbol{\mathcal{B}}_{4.3.1}+\boldsymbol{\mathcal{B}}_{4.3.3}\\
&=\left(\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0(\mathcal{M}_t^1+\mathcal{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)-\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0(\mathcal{M}_t^1+\mathcal{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\\
&\quad+\frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0(\mathcal{M}_t^1+\mathcal{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0(\mathcal{M}_t^1+\mathcal{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)\\
&\quad\times(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)\\
&\quad+\frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0(\mathcal{M}_t^1+\mathcal{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\varepsilon}_\tau,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0(\mathcal{M}_t^1+\mathcal{M}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\varepsilon}_\tau\right)\\
&\quad\times(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)+\boldsymbol{\Delta}_b+\boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2+\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0),
\end{aligned}\tag{G.169}$$

where $\boldsymbol{\Delta}_b$ is of the same order as $\boldsymbol{\Delta}_a$. Therefore, combining (G.140), (G.165), (G.166) and (G.169) gives

$$\boldsymbol{\mathcal{B}}_4=\left(\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)-\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^0\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau$$

$$
\begin{aligned}
&+ \left(\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_\tau\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\\
&- \frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\\
&+ \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)\\
&+ \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\varepsilon}_\tau,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\varepsilon}_\tau\right)(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)\\
&+ \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{f}_\tau^0\right)\\
&\times(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)\\
&+ \frac{1}{nT}\frac{1}{nT}\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_1\boldsymbol{\varepsilon}_\tau,...,\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{G}_Q\boldsymbol{\varepsilon}_\tau\right)\\
&\times(\hat{\boldsymbol{\rho}}-\boldsymbol{\rho}^0)+\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{Z}_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t+\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t\\
&+ \boldsymbol{\Delta}_a+\boldsymbol{\Delta}_b+\boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2+\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0). \qquad \text{(G.170)}
\end{aligned}
$$

Recall from Appendix A $\boldsymbol{Z}_t^* \coloneqq (\boldsymbol{W}_1\boldsymbol{y}_t,...,\boldsymbol{W}_Q\boldsymbol{y}_t,\boldsymbol{X}_t)$ and let $\boldsymbol{\Delta}_1 \coloneqq \boldsymbol{\Delta}_a+\boldsymbol{\Delta}_b$. Then, since $\boldsymbol{Z}_t+\boldsymbol{\mathcal{M}}_t^1+\boldsymbol{\mathcal{M}}_t^2=\boldsymbol{Z}_t^*$, (G.170) can be significantly simplified by gathering together many of the terms, resulting in

$$
\begin{aligned}
\boldsymbol{\mathcal{B}}_4 &= \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_t-\frac{1}{nT}\frac{1}{T}\sum_{t=1}^{T}\sum_{\tau=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}_\tau\varpi_{\tau t}^0\\
&+\frac{1}{nT}\frac{1}{T}\begin{pmatrix}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{W}_1\boldsymbol{y}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{W}_1\boldsymbol{y}_\tau & \cdots & \sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}(\boldsymbol{W}_1\boldsymbol{y}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{x}_{K\tau}\\ \vdots & \ddots & \vdots\\ \sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{x}_{Kt}'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{W}_1\boldsymbol{y}_\tau & \cdots & \sum_{t=1}^{T}\sum_{\tau=1}^{T}\varpi_{\tau t}^{0}\boldsymbol{x}_{Kt}'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{x}_{K\tau}\end{pmatrix}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\\
&+\boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2+\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\\
&= \frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*'}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon})+\frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*'}(\boldsymbol{P}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\mathcal{Z}}^*(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\\
&+\boldsymbol{O}_P\left(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2+\frac{P}{\sqrt{\min\{n,T\}}}\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0). \qquad \text{(G.171)}
\end{aligned}
$$

This completes the proof. $\square$

Proof of Lemma F.3(v). Recall that $\hat{\sigma}^2 \coloneqq \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{y}_t - \boldsymbol{X}_t\hat{\boldsymbol{\beta}})'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{y}_t - \boldsymbol{X}_t\hat{\boldsymbol{\beta}})$. Using Lemma A.2(i) and the true DGP, one obtains $\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{y}_t - \boldsymbol{X}_t\hat{\boldsymbol{\beta}} = \boldsymbol{Z}_t^*(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t$. Thus the expression for $\hat{\sigma}^2$ can be expanded to give

$$
\begin{aligned}
\hat{\sigma}^2 &= (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\left(\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t^*(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)\right) \\
&\quad + (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\left(\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)\right) + \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) \\
&\eqqcolon (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\boldsymbol{l}_1 + (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\boldsymbol{l}_2 + l_3,
\end{aligned}
\tag{G.172}
$$

where $\boldsymbol{Z}_t^* \coloneqq (\boldsymbol{W}_1\boldsymbol{y}_t, ..., \boldsymbol{W}_q\boldsymbol{y}_t, \boldsymbol{X}_t)$. Consider the term $(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\boldsymbol{l}_2$ in equation (G.172) first. Notice that $\boldsymbol{l}_2$ is equal to $\boldsymbol{\mathcal{B}}_4$ in equation (G.140). Hence, Lemma F.3(iv) can be applied to give

$$
\begin{aligned}
(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\boldsymbol{l}_2 = (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\Bigg(&\frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon}) \\
&+ \left(\boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
&+ \mathcal{H}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + \boldsymbol{\Delta}_1\Bigg).
\end{aligned}
\tag{G.173}
$$

An analogous expansion of term l_3 in equation (G.172) yields

$$
\begin{aligned}
l_3 &= \mathrm{tr}((\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} + \boldsymbol{\varepsilon})'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) + \mathcal{H}^*(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
&\quad + \Delta_2 + \left(\boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
&= \mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) + \mathcal{H}^*(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
&\quad + \Delta_2 + \left(\boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0),
\end{aligned}
\tag{G.174}
$$

since $\mathrm{tr}((\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) = 0$, and where $\mathcal{H}^* \coloneqq \frac{1}{nT}\mathrm{vec}(\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} + \boldsymbol{\varepsilon})'(\boldsymbol{P}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\mathcal{Z}}$ and Δ_2 is a term of the same order as $\boldsymbol{\Delta}_1$. Combining (G.172), (G.173) and (G.174),

$$
\hat{\sigma}^2 = (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})'\left(\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t^* - \mathcal{H}\right)(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})
$$

$$
\begin{aligned}
&+\frac{2}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon})+\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})+\Delta_2\\
&+\left(\boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0),
\end{aligned}
\tag{G.175}
$$

where the second equality follows by rearranging and noticing that $\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0+\boldsymbol{\varepsilon}_t)-(\boldsymbol{\mathcal{H}}^*)'=\frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon})$. Using Lemma F.3(iii),

$$
\begin{aligned}
&\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t^*)'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Z}_t^*-\boldsymbol{\mathcal{H}}\\
&=\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\boldsymbol{\mathcal{Z}}+\sigma_0^2\begin{pmatrix}\begin{pmatrix}\frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1'\boldsymbol{G}_1) & ... & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_1'\boldsymbol{G}_Q)\\ \vdots & \ddots & \vdots\\ \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q'\boldsymbol{G}_1) & ... & \frac{1}{n}\mathrm{tr}(\boldsymbol{G}_Q'\boldsymbol{G}_Q)\end{pmatrix} & \mathbf{0}_{Q\times K}\\ \mathbf{0}_{K\times Q} & \mathbf{0}_{K\times K}\end{pmatrix}\\
&\quad+\ \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\\
&=:\boldsymbol{\mathcal{K}}+\boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2)+\boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right).
\end{aligned}
\tag{G.176}
$$

In addition, using Lemma F.2(vii),

$$
\begin{aligned}
&\frac{1}{nT}\boldsymbol{\mathcal{Z}}^{*\prime}(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon})\\
&=\frac{1}{nT}(\boldsymbol{\mathcal{Z}}^*)'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}})\mathrm{vec}(\boldsymbol{\varepsilon})+\boldsymbol{o}_P\left(\frac{1}{\sqrt{nT}}\right)\\
&=\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon})+\frac{1}{nT}\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots\\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \mathbf{0}_{K\times 1}\end{pmatrix}\\
&\quad+\frac{1}{nT}\begin{pmatrix}\mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \vdots\\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right)\\ \mathbf{0}_{K\times 1}\end{pmatrix}+\boldsymbol{o}_P\left(\frac{1}{\sqrt{nT}}\right)
\end{aligned}
$$

$$
=: \frac{1}{nT}\mathcal{\boldsymbol{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) + \frac{1}{nT}\begin{pmatrix} \mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix}
$$

$$
+ \frac{1}{nT}\begin{pmatrix} \mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} + \boldsymbol{o}_P\left(\frac{1}{\sqrt{nT}}\right), \tag{G.177}
$$

where all elements $\mathrm{tr}((\boldsymbol{G}_q\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})$ are zero since $\boldsymbol{M}_{\boldsymbol{F}^0}\boldsymbol{F}^0 = \boldsymbol{0}_{T\times R}$. Combining (G.175), (G.176) and (G.177) gives the result

$$
\begin{aligned}
\hat{\sigma}^2 = {} & (\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})'\left(\boldsymbol{\mathcal{K}} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}) \\
& + 2(\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}})'\left(\frac{1}{nT}\mathcal{\boldsymbol{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\mathrm{vec}(\boldsymbol{\varepsilon}) + \frac{1}{nT}\begin{pmatrix} \mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} + \boldsymbol{o}_P\left(\frac{1}{\sqrt{nT}}\right)\right) \\
& + \mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) + \Delta_2.
\end{aligned} \tag{G.178}
$$

Lastly, using Lemma F.3(vi)

$$
\frac{1}{nT}\begin{pmatrix} \mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} = \begin{pmatrix} \frac{\sigma_0^2}{n}\mathrm{tr}\left(\boldsymbol{G}_1\right) \\ \vdots \\ \frac{\sigma_0^2}{n}\mathrm{tr}\left(\boldsymbol{G}_Q\right) \\ \boldsymbol{0}_{K\times 1}\end{pmatrix} + \boldsymbol{O}_P\left(\sqrt{\frac{Q}{nT}}\right) \tag{G.179}
$$

and

$$
\begin{aligned}
& \frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) \\
= & \frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) + \frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0}-\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) \\
= & \frac{1}{nT}\mathrm{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) + O_P\left(\frac{\sqrt{Q}||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2}{\min\{n,T\}}\right) + O_P\left(\frac{1}{\min\{n^{1.5},T^{1.5}\}}\right),
\end{aligned} \tag{G.180}
$$

since $|\text{tr}(\boldsymbol{\varepsilon}'(\boldsymbol{P}_{\boldsymbol{\Lambda}^0} - \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}})\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0})| \leq 2R||\boldsymbol{\varepsilon}||_2^2||\boldsymbol{P}_{\boldsymbol{\Lambda}^0} - \boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}||_2||\boldsymbol{M}_{\boldsymbol{F}^0}||_2$. Combining (G.178), (G.179) and (G.180), and ignoring dominated terms gives the final result

$$\hat{\sigma}^2 = (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \left(\boldsymbol{\mathcal{K}} + \boldsymbol{O}_P(\sqrt{Q}P||\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}||_2) + \boldsymbol{O}_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right)\right)(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})$$
$$+ 2(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \left(\begin{pmatrix} \frac{\sigma_0^2}{n}\text{tr}\,(\boldsymbol{G}_1) \\ \vdots \\ \frac{\sigma_0^2}{n}\text{tr}\,(\boldsymbol{G}_Q) \\ \boldsymbol{0}_{K\times 1} \end{pmatrix} + \boldsymbol{O}_P\left(\sqrt{\frac{Q}{nT}}\right)\right) + \frac{1}{nT}\text{tr}(\boldsymbol{\varepsilon}'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}) + \Delta_2.$$

□

Proof of Lemma F.3(vi). First,

$$\frac{1}{\sqrt{nT}}\boldsymbol{b}_4 = \frac{1}{\sqrt{nT}}\begin{pmatrix} \text{tr}\,((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}) \\ \vdots \\ \text{tr}\,((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}) \end{pmatrix} - \frac{1}{\sqrt{nT}}\begin{pmatrix} \text{tr}\,((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}) \\ \vdots \\ \text{tr}\,((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}) \end{pmatrix}$$
$$- \frac{1}{\sqrt{nT}}\begin{pmatrix} \text{tr}\,((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \\ \vdots \\ \text{tr}\,((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \end{pmatrix} + \frac{1}{\sqrt{nT}}\begin{pmatrix} \text{tr}\,((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \\ \vdots \\ \text{tr}\,((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}) \end{pmatrix}$$
$$=: \frac{1}{\sqrt{nT}}(\boldsymbol{l}_1 + ... + \boldsymbol{l}_4). \tag{G.181}$$

Now,

$$\boldsymbol{l}_1 = \begin{pmatrix} \text{tr}\,((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}) - T\sigma_0^2\text{tr}(\boldsymbol{G}_1) \\ \vdots \\ \text{tr}\,((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}) - T\sigma_0^2\text{tr}(\boldsymbol{G}_Q) \end{pmatrix} + \begin{pmatrix} T\sigma_0^2\text{tr}(\boldsymbol{G}_1) \\ \vdots \\ T\sigma_0^2\text{tr}(\boldsymbol{G}_Q) \end{pmatrix}$$
$$=: \boldsymbol{l}_{1.1} + \begin{pmatrix} T\sigma_0^2\text{tr}(\boldsymbol{G}_1) \\ \vdots \\ T\sigma_0^2\text{tr}(\boldsymbol{G}_Q) \end{pmatrix}. \tag{G.182}$$

For the term $\boldsymbol{l}_{1.1}$,

$$\mathbb{E}\left[||\boldsymbol{l}_{1.1}||_2^2\right] = \mathbb{E}\left[\sum_{q=1}^{Q}\left(\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{G}_q'\boldsymbol{\varepsilon}_t - \sigma_0^2\text{tr}(\boldsymbol{G}_q)\right)^2\right]$$
$$= \sum_{q=1}^{Q} T\left((\mathcal{M}_4^{\varepsilon} - 3\sigma_0^4)\sum_{i=1}^{n}(G_q)_{ii}^2 + \sigma_0^4(\text{tr}(\boldsymbol{G}_q\boldsymbol{G}_q) + \text{tr}(\boldsymbol{G}_q'\boldsymbol{G}_q))\right)$$

$$= O(QnT) \tag{G.183}$$

using Lemma A.3 in Yu et al. (2008). Therefore $||\boldsymbol{l}_{1.1}||_2 = O_P(\sqrt{QnT})$. Next,

$$||\boldsymbol{l}_2||_2^2 = \sum_{q=1}^{Q} \mathrm{tr}\left((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\right)^2 \leq (R^0)^2 \frac{1}{n^2} \sum_{q=1}^{Q} ||(\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{\Lambda}^0||_2^2 \left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2^2 ||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}||_2^2$$
$$= O_P(QT^2), \tag{G.184}$$

where Lemma F.1(ii) has been applied, with $\boldsymbol{\Lambda}^0$ replaced by $\boldsymbol{G}_q'\boldsymbol{\Lambda}^0$, to establish that $\sum_{q=1}^{Q} ||(\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{\Lambda}^0||_2^2 = O_P(QnT)$. Thus $||\boldsymbol{l}_2||_2 = O_P(\sqrt{Q}T)$. Next for term $\boldsymbol{l}_3$,

$$||\boldsymbol{l}_3||_2^2 = \sum_{q=1}^{Q} \mathrm{tr}\left((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)^2 = \frac{1}{T^2}\sum_{q=1}^{Q} \mathrm{tr}\left(\boldsymbol{G}_q'\boldsymbol{\varepsilon}\boldsymbol{F}^0\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0\prime}\boldsymbol{\varepsilon}'\right)^2$$
$$= \frac{1}{T^2}(R^0)\sum_{q=1}^{Q} ||\boldsymbol{G}_q||_2^2 ||\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2^4 \left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2^2 = O_P(Qn^2) \tag{G.185}$$

using Lemma F.1(ii). Thus $||\boldsymbol{l}_3||_2 = O_P(\sqrt{Q}n)$. Finally,

$$||\boldsymbol{l}_4||_2^2 = \sum_{q=1}^{Q} \mathrm{tr}\left((\boldsymbol{G}_q\boldsymbol{\varepsilon})'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{P}_{\boldsymbol{F}^0}\right)^2$$
$$= \frac{1}{n^2T^2}\sum_{q=1}^{Q} \mathrm{tr}\left(\boldsymbol{\varepsilon}'\boldsymbol{G}_q'\boldsymbol{\Lambda}^0\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\boldsymbol{F}^{0\prime}\right)^2$$
$$\leq \frac{1}{n^2T^2}(R^0)\sum_{q=1}^{Q} ||\boldsymbol{\varepsilon}'\boldsymbol{G}_q'\boldsymbol{\Lambda}^0||_2^2 \left|\left|\left(\frac{1}{n}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^{-1}\right|\right|_2^2 ||\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0||_2^2 \left|\left|\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{-1}\right|\right|_2^2 ||\boldsymbol{F}^0||_2^2$$
$$= O_P(QT), \tag{G.186}$$

using Lemmas A.2(iii), F.1(ii) and F.1(ii). Hence $||\boldsymbol{l}_4||_2 = O_P(\sqrt{QT})$. Combining all these results

$$\frac{1}{\sqrt{nT}}\begin{pmatrix} \mathrm{tr}\left((\boldsymbol{G}_1\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \\ \vdots \\ \mathrm{tr}\left((\boldsymbol{G}_Q\boldsymbol{\varepsilon})'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\boldsymbol{M}_{\boldsymbol{F}^0}\right) \end{pmatrix} = \frac{1}{\sqrt{nT}}\begin{pmatrix} T\sigma_0^2\mathrm{tr}(\boldsymbol{G}_1) \\ \vdots \\ T\sigma_0^2\mathrm{tr}(\boldsymbol{G}_Q) \end{pmatrix} + \boldsymbol{O}_P(\sqrt{Q}), \tag{G.187}$$

since, by Assumption 6.4, $\frac{T}{n} \to c$. This completes the proof. □

Proof of Lemma F.4(i).

$$\sum_{p=1}^{P}\left|\left|\sum_{k=1}^{K}(\hat{\beta}_k^0 - \beta_k)(\boldsymbol{G}_p(\hat{\boldsymbol{\rho}}) - \boldsymbol{G}_p)\boldsymbol{\mathcal{X}}_k\right|\right|_F^2$$

$$
\begin{aligned}
&= \sum_{q=1}^{Q}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\sum_{k=1}^{K}(\hat{\beta}_k^0-\beta_k)((\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})-\boldsymbol{G}_q)\boldsymbol{\mathcal{X}}_k)_{it}\right)^2 \\
&= \sum_{q=1}^{Q}\sum_{t=1}^{T}||(\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})-\boldsymbol{G}_q)\boldsymbol{X}_t(\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}})||_2^2 \\
&\leq ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})-\boldsymbol{G}_q||_2^2\sum_{t=1}^{T}||\boldsymbol{X}_t||_2^2 \\
&\leq ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})-\boldsymbol{G}_q||_2^2\sum_{t=1}^{T}||\boldsymbol{X}_t||_F^2 \\
&= ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})-\boldsymbol{G}_q||_2^2\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_F^2 \\
&= ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2\sum_{q=1}^{Q}||\boldsymbol{W}_q\boldsymbol{S}^{-1}(\hat{\boldsymbol{\rho}})(\boldsymbol{I}_n-\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1})||_2^2\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_F^2 \\
&\leq ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2||\boldsymbol{S}^{-1}(\hat{\boldsymbol{\rho}})||_2^2||\boldsymbol{I}_n-\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}||_2^2 Q\max_{1\leq q\leq Q}\{||\boldsymbol{W}_q||_2^2\}\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_F^2 \\
&= O_P(||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^4)O_P(Q^2KnT) \qquad (G.188)
\end{aligned}
$$

using Lemmas A.2(ii) and A.2(viii). □

Proof of Lemma F.4(ii).

$$
\begin{aligned}
\sum_{p=1}^{P}\left|\left|\sum_{k=1}^{K}(\hat{\beta}_k^0-\beta_k)\boldsymbol{G}_p\boldsymbol{\mathcal{X}}_k\right|\right|_F^2 &= \sum_{q=1}^{Q}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\sum_{k=1}^{K}(\hat{\beta}_k^0-\beta_k)(\boldsymbol{G}_q\boldsymbol{\mathcal{X}}_k)_{it}\right)^2 \\
&= \sum_{q=1}^{Q}\sum_{t=1}^{T}||\boldsymbol{G}_q\boldsymbol{X}_t(\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}})||_2^2 \\
&\leq ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q||_2^2\sum_{t=1}^{T}||\boldsymbol{X}_t||_2^2 \\
&\leq ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q||_2^2\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_F^2 \\
&\leq ||\boldsymbol{\beta}^0-\hat{\boldsymbol{\beta}}||_2^2 Q\max_{1\leq q\leq Q}\{||\boldsymbol{G}_q||_2^2\}\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_F^2 \\
&= O_P(||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2)O_P(QKnT) \qquad (G.189)
\end{aligned}
$$

using Lemmas A.2(ii) and A.2(viii). $\square$

Proof of Lemma F.4(iii).

$$\begin{aligned}
\sum_{p=1}^{P}\left|\left|\sum_{k=1}^{K}\beta_k^0(\boldsymbol{G}_p-\boldsymbol{G}_p(\hat{\boldsymbol{\rho}}))\boldsymbol{\mathcal{X}}_k\right|\right|_F^2 &= \sum_{q=1}^{Q}\left|\left|\sum_{k=1}^{K}\beta_k^0(\boldsymbol{G}_q-\boldsymbol{G}_q(\hat{\boldsymbol{\rho}}))\boldsymbol{\mathcal{X}}_k\right|\right|_F^2 \\
&\leq \sum_{q=1}^{Q}\sum_{t=1}^{T}||(\boldsymbol{G}_q-\boldsymbol{G}_q(\hat{\boldsymbol{\rho}}))\boldsymbol{X}_t\boldsymbol{\beta}^0||_2^2 \\
&= ||\boldsymbol{\beta}^0||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q-\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2^2\sum_{t=1}^{T}||\boldsymbol{X}_t||_2^2 \\
&\leq ||\boldsymbol{\beta}^0||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q-\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2^2\sum_{t=1}^{T}||\boldsymbol{X}_t||_F^2 \\
&= ||\boldsymbol{\beta}^0||_2^2\sum_{q=1}^{Q}||\boldsymbol{W}_q\boldsymbol{S}^{-1}-\boldsymbol{W}_q\boldsymbol{S}^{-1}(\hat{\boldsymbol{\rho}})||_2^2\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_2^2 \\
&= ||\boldsymbol{\beta}^0||_2^2 Q\max_{1\leq q\leq Q}\{||\boldsymbol{W}_q||_2^2\}||\boldsymbol{S}(\hat{\boldsymbol{\rho}})||_2^2||\boldsymbol{S}(\hat{\boldsymbol{\rho}})\boldsymbol{S}^{-1}-\boldsymbol{I}_n||_2^2\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_2^2 \\
&= O_P(||\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}||_2^2)O_P(Q^2KnT) \qquad \text{(G.190)}
\end{aligned}$$

using Lemmas A.2(v) and A.2(viii). $\square$

Proof of Lemma F.4(iv).

$$\begin{aligned}
\sum_{p=1}^{P}\left|\left|\sum_{k=1}^{K}(\hat{\beta}_k-\beta_k^0)\boldsymbol{G}(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k\right|\right|_F^2 &= \sum_{q=1}^{Q}\left|\left|\sum_{k=1}^{K}(\hat{\beta}_k-\beta_k^0)\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k\right|\right|_F^2 \\
&= \sum_{q=1}^{Q}\sum_{t=1}^{T}||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})\boldsymbol{X}_t(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)||_2^2 \\
&\leq ||\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2\sum_{t=1}^{T}||\boldsymbol{X}_t||_2^2 \\
&\leq ||\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0||_2^2 Q\max_{1\leq q\leq Q}\{||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2\}\sum_{t=1}^{T}||\boldsymbol{X}_t||_F^2 \\
&\leq ||\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0||_2^2 Q\max_{1\leq q\leq Q}\{||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2\}\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}_k||_F^2 \\
&= O_P(||\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0||_2^2)O_P(QKnT) \qquad \text{(G.191)}
\end{aligned}$$

using Lemmas A.2(v) and A.2(viii). □

Proof of Lemma F.4(v).

$$
\begin{aligned}
\sum_{p=1}^{P}\left|\left|\sum_{k=1}^{K}\beta_k^0\boldsymbol{G}(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k\right|\right|_F^2 &= \sum_{q=1}^{Q}\left|\left|\sum_{k=1}^{K}\beta_k^0\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})\boldsymbol{\mathcal{X}}_k\right|\right|_F^2 \\
&= \sum_{q=1}^{Q}\sum_{t=1}^{T}||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})\boldsymbol{X}_t\boldsymbol{\beta}^0||_2^2 \\
&\leq ||\boldsymbol{\beta}^0||_2^2\sum_{q=1}^{Q}||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2^2\sum_{t=1}^{T}||\boldsymbol{X}_t||_2^2 \\
&\leq ||\boldsymbol{\beta}^0||_2^2 Q\max_{1\leq q\leq Q}\{||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2^2\}\sum_{t=1}^{T}||\boldsymbol{X}_t||_F^2 \\
&= ||\boldsymbol{\beta}^0||_2^2 Q\max_{1\leq q\leq Q}\{||\boldsymbol{G}_q(\hat{\boldsymbol{\rho}})||_2^2\}\sum_{k=1}^{K}||\boldsymbol{\mathcal{X}}||_F^2 \\
&= O_P(QKnT) \qquad \text{(G.192)}
\end{aligned}
$$

using Lemmas A.2(v) and A.2(viii). □

H Proof of Proposition ID

What follows is analogous to the proofs provided for Theorem 2.1 in Moon and Weidner (2015) and Proposition 1 in Shi and Lee (2017). The expected unpenalised likelihood, evaluated at some $\boldsymbol{\theta}, \boldsymbol{\Lambda}$ and $\boldsymbol{F}$, and with σ^2 concentrated out, is denoted $\mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \boldsymbol{F})$. Dropping the constant this is given by

$$
\begin{aligned}
\mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \boldsymbol{F}) = &\frac{1}{n}\log(\det(\boldsymbol{S}(\boldsymbol{\rho}))) \\
&- \frac{1}{2}\log\left(\mathbb{E}\left[\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{y}_t - \boldsymbol{X}_t\boldsymbol{\beta} - \boldsymbol{\Lambda}\boldsymbol{f}_t)'(\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{y}_t - \boldsymbol{X}_t\boldsymbol{\beta} - \boldsymbol{\Lambda}\boldsymbol{f}_t)\right]\right).
\end{aligned}
\tag{H.1}
$$

Substituting the true DGP $\boldsymbol{y}_t = \boldsymbol{S}^{-1}(\boldsymbol{X}_t\boldsymbol{\beta}^0 + \boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)$ into (H.1) and applying Lemma A.2(i) results in

$$
\mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \boldsymbol{F}) = \frac{1}{n}\log(\det(\boldsymbol{S}(\boldsymbol{\rho})))
$$

$$-\frac{1}{2}\log\Big(\mathbb{E}\Big[\frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\boldsymbol{\theta})+\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0+\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t-\boldsymbol{\Lambda}\boldsymbol{f}_t)'$$

$$\times(\boldsymbol{Z}_t(\boldsymbol{\theta}^0-\boldsymbol{\theta})+\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0+\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}_t-\boldsymbol{\Lambda}\boldsymbol{f}_t)\Big]\Big). \tag{H.2}$$

Now, to begin, it is shown that for any $(\boldsymbol{\theta},\boldsymbol{\Lambda}\boldsymbol{F})\neq(\boldsymbol{\theta}^0,\boldsymbol{\Lambda}^0\boldsymbol{F}^0)$, $\mathbb{L}(\boldsymbol{\theta},\boldsymbol{\Lambda},\boldsymbol{F})<\mathbb{L}(\boldsymbol{\theta}^0,\boldsymbol{\Lambda}^0,\boldsymbol{F}^0)$. First of all,

$$\begin{aligned}\mathbb{L}(\boldsymbol{\theta}^0,\boldsymbol{\Lambda}^0,\boldsymbol{F}^0)&=\frac{1}{n}\log(\det(\boldsymbol{S}))-\frac{1}{2}\log\Big(\mathbb{E}\Big[\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t\Big]\Big),\\&=\frac{1}{n}\log(\det(\boldsymbol{S}))-\frac{1}{2}\log(\sigma_0^2),\end{aligned}\tag{H.3}$$

where the second line follows by Assumption 1.1. Next, using Assumption 1, and concentrating out $\boldsymbol{F}$ and $\boldsymbol{\Lambda}^0$, gives the inequality

$$\begin{aligned}\mathbb{L}(\boldsymbol{\theta},\boldsymbol{\Lambda},\boldsymbol{F})\leq\frac{1}{n}\log(\det(\boldsymbol{S}(\boldsymbol{\rho})))-\frac{1}{2}\log\Bigg(&\frac{1}{nT}\mathbb{E}\big[\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon})\big]\\&+\mathbb{E}\Bigg[\mathrm{tr}\Bigg(\boldsymbol{M}_{\boldsymbol{F}^0}\Bigg(\sum_{p=1}^{P}(\theta_p^0-\theta_p)\mathcal{Z}_p\Bigg)'\boldsymbol{M}_{\boldsymbol{\Lambda}}\Bigg(\sum_{p=1}^{P}(\theta_p^0-\theta_p)\mathcal{Z}_p\Bigg)\Bigg)\Bigg]\Bigg).\end{aligned}\tag{H.4}$$

By Lemma 9 in Yu et al. (2008), $\frac{1}{nT}\mathbb{E}[\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon})'(\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}\boldsymbol{\varepsilon}))]=\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})$. Applying this, and then rearranging (H.4) gives

$$\begin{aligned}\mathbb{L}(\boldsymbol{\theta},\boldsymbol{\Lambda},\boldsymbol{F})\leq&\frac{1}{n}\log(\det(\boldsymbol{S}(\boldsymbol{\rho})))\\&-\frac{1}{2}\log\Big(\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})+\mathbb{E}\Big[\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}})\boldsymbol{\mathcal{Z}}\Big]\Big).\end{aligned}\tag{H.5}$$

For simplicity denote $\mathbb{E}\Big[\frac{1}{nT}\boldsymbol{\mathcal{Z}}'(\boldsymbol{M}_{\boldsymbol{F}^0}\otimes\boldsymbol{M}_{\boldsymbol{\Lambda}})\boldsymbol{\mathcal{Z}}\Big]$ by $\mathbb{M}$. Now $\boldsymbol{\theta}^0$ is a unique global maximiser of the unpenalised expected likelihood for any $\boldsymbol{\Lambda},\boldsymbol{F}$, if for any $\boldsymbol{\theta}\neq\boldsymbol{\theta}^0$, $\mathbb{L}(\boldsymbol{\theta},\boldsymbol{\Lambda},\boldsymbol{F})<\mathbb{L}(\boldsymbol{\theta}^0,\boldsymbol{\Lambda},\boldsymbol{F})$. Using (H.3) and (H.5), this inequality holds when

$$\begin{aligned}&\frac{1}{n}\log(\det(\boldsymbol{S}(\boldsymbol{\rho})))-\frac{1}{2}\log\Big(\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})+(\boldsymbol{\theta}^0-\boldsymbol{\theta})'\mathbb{M}(\boldsymbol{\theta}^0-\boldsymbol{\theta})\Big)\\&\qquad<\frac{1}{n}\log(\det(\boldsymbol{S}))-\frac{1}{2}\log(\sigma_0^2).\end{aligned}\tag{H.6}$$

Note that $-\frac{1}{n}\log(\det(\boldsymbol{S}(\boldsymbol{\rho})))+\frac{1}{n}\log(\det(\boldsymbol{S}))-\frac{1}{2}\log(\sigma_0^2)=-\frac{1}{2}\log(\sigma_0^2\det((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})^{\frac{1}{n}})$. Therefore (H.6) becomes

$$-\frac{1}{2}\log\Big(\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})+(\boldsymbol{\theta}^0-\boldsymbol{\theta})'\mathbb{M}(\boldsymbol{\theta}^0-\boldsymbol{\theta})\Big)$$

$$< -\frac{1}{2}\log(\sigma_0^2 \det((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})^{\frac{1}{n}}). \tag{H.7}$$

Multiplying by $-\frac{1}{2}$ and then raising both sides as a power of e yields the condition

$$\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}) + (\boldsymbol{\theta}^0 - \boldsymbol{\theta})'\mathbb{M}(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) > \sigma_0^2 \det((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})^{\frac{1}{n}}. \tag{H.8}$$

The matrix $\mathbb{M}$ is positive definite by Assumption ID.2 and, moreover, by Lemma A.1, $\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}) \geq \sigma_0^2 \det((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})^{\frac{1}{n}}$. Hence,

$$\frac{\sigma_0^2}{n}\mathrm{tr}((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1}) - \sigma_0^2 \det((\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})'\boldsymbol{S}(\boldsymbol{\rho})\boldsymbol{S}^{-1})^{\frac{1}{n}} + (\boldsymbol{\theta}^0 - \boldsymbol{\theta})'\mathbb{M}(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) > 0, \tag{H.9}$$

for any $\boldsymbol{\theta} \neq \boldsymbol{\theta}^0$ and the expected likelihood must be uniquely maximised at $\boldsymbol{\theta}^0$ for any $\boldsymbol{\Lambda}, \boldsymbol{F}$. Now, evaluated at $\boldsymbol{\theta}^0$ and with $\boldsymbol{\Lambda}\boldsymbol{F}' \neq \boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}$,

$$\begin{aligned}\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}, \boldsymbol{F}) &\leq \frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\log\left(\mathbb{E}\Big[\frac{1}{nT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t + \frac{1}{nT}\sum_{t=1}^{T}(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 - \boldsymbol{\Lambda}\boldsymbol{f}_t)'(\boldsymbol{\Lambda}^0\boldsymbol{f}_t^0 - \boldsymbol{\Lambda}\boldsymbol{f}_t)\Big]\right)\\ &= \frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\log\left(\sigma_0^2 + \mathbb{E}\Big[\frac{1}{nT}\mathrm{tr}((\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} - \boldsymbol{\Lambda}\boldsymbol{F}')'(\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} - \boldsymbol{\Lambda}\boldsymbol{F}'))\Big]\right).\end{aligned} \tag{H.10}$$

The trace term in (H.10) can be equivalently written as $\mathrm{vec}(\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} - \boldsymbol{\Lambda}\boldsymbol{F}')'(\boldsymbol{I}_T \otimes \boldsymbol{I}_n)\mathrm{vec}(\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} - \boldsymbol{\Lambda}\boldsymbol{F}')$. Since the matrix $(\boldsymbol{I}_T \otimes \boldsymbol{I}_n)$ is positive definite, this term is strictly positive as long as $\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} \neq \boldsymbol{\Lambda}\boldsymbol{F}'$ and therefore

$$\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}, \boldsymbol{F}) < \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \boldsymbol{F}^0) = \frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\log\left(\sigma_0^2\right) \tag{H.11}$$

and the expected likelihood is maximised where $\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime} = \boldsymbol{\Lambda}\boldsymbol{F}'$. The inequality (H.11) implies extremum identification of $\boldsymbol{\theta}^0$ and the product $\boldsymbol{\Lambda}^0\boldsymbol{F}^{0\prime}$. In a quasi-likelihood setting, this is sufficient for identification of these parameters. Identification of σ_0^2 is then straightforward to show since, omitting the constant, $\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \boldsymbol{F}^0, \sigma_0^2) = -\frac{1}{2}\log(\sigma_0^2) + \frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}$ and $\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \boldsymbol{F}^0, \sigma^2) = -\frac{1}{2}\log(\sigma^2) + \frac{1}{n}\log(\det(\boldsymbol{S})) - \frac{1}{2}\frac{\sigma_0^2}{\sigma^2}$. Thus $\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \boldsymbol{F}^0, \sigma^2) < \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \boldsymbol{F}^0, \sigma_0^2)$ holds if

$$-\frac{1}{2}\log(\sigma^2) - \frac{1}{2}\frac{\sigma_0^2}{\sigma^2} < -\frac{1}{2}\log(\sigma_0^2) - \frac{1}{2} \tag{H.12}$$

or

$$\log\left(\frac{\sigma_0^2}{\sigma^2}\right) < \frac{\sigma_0^2}{\sigma^2} - 1. \tag{H.13}$$

Using $\log(x) < x - 1$ for $x > 1$ and $x \neq 1$, it is clear that σ_0^2 is also identified.

I Verifying Assumptions 1–8

In this appendix, where possible, the assumptions in the main text are verified for the Monte Carlo experiment detailed in Section 5.1.

2.2 The q-th weights matrix is constructed as if all the cross-sectional units were arrayed on a line and connected only to their immediate neighbours to the q-th degree. These weights matrices are then row normalised. Each row sum is 1 and for $Q \geq n/4$, which is satisfied in the Monte Carlo design, the maximum column sum of each matrix is 1.5. Hence each weights matrix is UB uniformly over q.

2.3 Invertibility of $\boldsymbol{S}(\boldsymbol{\rho})$: For simplicity, let $\boldsymbol{W}^* := \sum_{q=1}^{Q} \rho_q \boldsymbol{W}_q$. A sufficient condition for the invertibility of $\boldsymbol{S}(\boldsymbol{\rho}) = \boldsymbol{I}_n - \boldsymbol{W}^*$ is that $||\boldsymbol{W}^*|| < 1$ for some norm $||\cdot||$.[3] Since

$$||\boldsymbol{W}^*|| \leq \sum_{q=1}^{Q} |\rho_q| \max_{1 \leq q \leq Q} ||\boldsymbol{W}_q||, \tag{I.1}$$

$\boldsymbol{S}(\boldsymbol{\rho})$ will be invertible if $\sum_{q=1}^{Q} |\rho_q| < (\max_{1 \leq q \leq Q} ||\boldsymbol{W}_q||)^{-1}$. In particular, since all the $\boldsymbol{W}_q$ are row normalised (so that $||\boldsymbol{W}_q||_\infty = 1$), this condition reduces to $\sum_{q=1}^{Q} |\rho_q| < 1$, which is satisfied in the MC design.

$\boldsymbol{S}(\boldsymbol{\rho})$ is UB: Since $||\boldsymbol{S}(\boldsymbol{\rho})|| = ||\boldsymbol{I}_n - \boldsymbol{W}^*|| \leq ||\boldsymbol{I}_n|| + ||\boldsymbol{W}^*||$, $\boldsymbol{S}(\boldsymbol{\rho})$ is UB if $\boldsymbol{W}^*$ is UB. If $\sum_{q=1}^{Q} |\rho_q| < 1$, this follows from equation (I.1) and Assumption 2.2, which has been verified above.

$\boldsymbol{S}^{-1}(\boldsymbol{\rho})$ is UB: If $||\boldsymbol{W}^*|| < 1$ for some norm $||\cdot||$, then $\boldsymbol{S}^{-1}(\boldsymbol{\rho}) = \sum_{h=0}^{\infty} (\boldsymbol{W}^*)^h$ and therefore

$$||\boldsymbol{S}^{-1}(\boldsymbol{\rho})|| \leq \sum_{h=0}^{\infty} ||(\boldsymbol{W}^*)^h|| \leq \sum_{h=0}^{\infty} ||(\boldsymbol{W}^*)||^h. \tag{I.2}$$

Under the condition $\sum_{q=1}^{Q} |\rho_q| < 1$, $||\boldsymbol{W}^*||_\infty < 1$ (see above) and therefore $||\boldsymbol{S}^{-1}(\boldsymbol{\rho})||_\infty$ is bounded, by equation (I.2). For the columns sums, since the absolute column sums of each weights matrix are bounded by 1.5, by (I.1), $||\boldsymbol{W}^*||_1$ will be less than 1 where $\sum_{q=1}^{Q} |\rho_q| < 1/1.5$. Using (I.2), it is then straightforward to demonstrate that a sufficient condition for $\boldsymbol{S}^{-1}(\boldsymbol{\rho})$ to be uniformly bounded in absolute column sums is $\sum_{q=1}^{Q} |\rho_q| < 1/1.5$, which is satisfied in the Monte Carlo design.

[3] Easily verifiable by considering the Neumann series of $(\boldsymbol{I}_n - \boldsymbol{W})^{-1}$.

3.1, 5, 6.6 In simulations, $\gamma_1 = ... = \gamma_Q = \gamma_{Q+K^*+1} = ... = \gamma_P$ and $\gamma_{Q+1} = ...\gamma_{Q+K^*}$ are imposed, and the information criterion in described Section 4.1 is used to select the penalty parameter.

3.2 As an initial estimate, the unpenalised MLE is used. In Proposition 1 this is shown to be at least a_{nT}-consistent with $R \geq R^0$.

4.3 In the design, the number of parameters adheres to a sequence, indexed say $j = 1, 2, ...$, where $n = T$, such that $P = 8 + 4j$ and $n = T = 25 \times 2^{j-1}$. For example, $j = 1, P = 12, n = T = 25$, $j = 2, P = 16, n = T = 50$, $j = 3, P = 20, n = T = 100$, etc. Thus, $\lim_{j\to\infty} \frac{P(j)}{\min\{n(j),T(j)\}} = \lim_{j\to\infty} \frac{8+4j}{25\times 2^{j-1}} = 0$.

6.1 Considering the same sequences as in 4.3, note that $\lim_{j\to\infty} \frac{P^4(j)}{\min\{n(j),T(j)\}} = \lim_{j\to\infty} \frac{(8+4j)^4}{25\times 2^{j-1}} = 0$, whereby Assumption 6.1 is satisfied.

8.1 Let $\varrho_\rho = \varrho_\beta = 1/\min\{n^{1/4}, T^{1/4}\}$. Then with $a_{nT} = \sqrt{P}/\sqrt{\min\{n,T\}}$, and the number of parameters adhering to the sequence described previously in 6.1, the sequences $\min\{n^{1/4}, T^{1/4}\}/\sqrt{PQ} \to \infty$ and $P/\min\{n^{1/4}, T^{1/4}\} \to 0$.

J Additional Tables

J.1 Standard Normal Errors

Table 6: Bias of penalised estimator of nonzero coefficients ($R = 0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.0245	−0.0038	-	−0.1183	0.1471	-	−0.1634	−0.0596	-	−0.0014	0.0012
25	50	0.0029	−0.0028	-	−0.1133	0.1432	-	−0.1648	−0.0593	-	−0.0019	0.0019
	100	0.0049	−0.0055	-	−0.1389	0.0153	0.0205	−0.0817	−0.0500	-	−0.0010	0.0010
	25	0.0205	−0.0009	−0.0007	0.1158	0.1466	-	−0.1561	−0.0768	-	−0.0014	0.0014
50	50	0.0223	−0.0022	−0.0014	0.1144	0.1448	-	−0.1637	−0.0778	-	−0.0017	0.0017
	100	0.0056	−0.0030	−0.0038	0.1428	0.0101	0.0261	−0.0790	0.0514	-	−0.0008	0.0008
	25	−0.0208	−0.0038	−0.0012	0.1219	0.1425	-	−0.1483	−0.0693	−0.0775	−0.0020	0.0020
100	50	−0.0207	−0.0039	−0.0012	0.1216	0.1409	-	−0.1505	−0.0702	−0.0778	−0.0020	0.0020
	100	−0.0041	−0.0031	−0.0021	0.1538	0.0103	0.0261	−0.0581	−0.0366	−0.0448	−0.0008	0.0008

Table 7: Bias of bias corrected estimates of nonzero parameters ($R = 1$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.0131	0.0009	-	0.0583	0.0729	-	−0.0899	−0.0286	-	−0.0013	0.0012
25	50	0.0141	0.0003	-	0.0612	0.0760	-	−0.0940	−0.0328	-	−0.0011	0.0011
	100	0.0017	−0.0029	-	0.0849	0.0220	0.0268	−0.0380	−0.0233	-	−0.0005	0.0006
	25	0.0122	−0.0011	−0.0010	0.0635	0.0790	-	−0.0905	−0.0411	-	−0.0007	0.0008
50	50	0.0131	−0.0021	−0.0008	0.0660	0.0817	-	−0.0945	−0.0438	-	−0.0010	0.0010
	100	0.0020	−0.0015	−0.0019	0.0958	0.0326	0.0411	−0.0344	−0.0228	-	−0.0005	0.0008
	25	0.0115	−0.0022	0.0007	0.0671	0.0774	-	−0.0830	−0.0384	−0.0432	−0.0012	0.0012
100	50	0.0119	−0.0023	0.0007	0.0704	0.0813	-	−0.0849	−0.0404	−0.0445	−0.0012	0.0011
	100	0.0014	−0.0013	−0.0012	0.1056	0.0403	0.0479	−0.0236	−0.0162	−0.0185	−0.0006	0.0006

Table 8: Coverage of nonzero parameter estimates ($R = 1$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.631	0.885	-	0.465	0.337	-	0.410	0.817	-	0.914	0.908
25	50	0.393	0.882	-	0.200	0.095	-	0.141	0.715	-	0.893	0.899
	100	0.969	0.664	-	0.006	0.776	0.718	0.403	0.608	-	0.886	0.880
	25	0.457	0.907	0.894	0.159	0.073	-	0.172	0.588	-	0.902	0.904
50	50	0.179	0.847	0.865	0.020	0.003	-	0.031	0.374	-	0.897	0.899
	100	0.615	0.679	0.639	0.000	0.709	0.486	0.283	0.471	-	0.870	0.854
	25	0.325	0.865	0.921	0.021	0.009	-	0.109	0.493	0.408	0.880	0.880
100	50	0.084	0.820	0.868	0.000	0.000	-	0.012	0.251	0.164	0.867	0.863
	100	0.636	0.655	0.628	0.000	0.602	0.248	0.318	0.470	0.416	0.786	0.796

Table 9: Bias of bias corrected estimates of nonzero parameters ($R = 5$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.0001	−0.0006	-	0.0011	−0.0019	-	−0.0022	0.0037	-	−0.0005	0.0004
25	50	0	−0.0003	-	0.0001	−0.0012	-	−0.0012	0.0029	-	−0.0002	0.0002
	100	0	−0.0001	-	0.0002	−0.0007	0.0007	−0.0009 E	0.0014	-	−0.0001	0.0001
	25	0.0001	−0.0001	−0.0001	0.0006	−0.0007	-	0	0.0007	-	−0.0002	0.0002
50	50	0.0002	−0.0003	0	0.0005	−0.0012	-	−0.0008	0.0019	-	−0.0002	0.0002
	100	0	−0.0001	0	−0.0001	−0.0005	0.0005	0.0002	0.0005	-	−0.0002	0.0002
	25	−0.0001	−0.0002	0.0002	0.0003	−0.0011	-	−0.0004	0.0022	−0.0006	−0.0003	0.0003
100	50	0	−0.0001	0	0.0004	−0.0003	-	0.0002	0.0006	−0.0006	−0.0002	0.0003
	100	0	−0.0001	0	0.0001	−0.0001	0.0002	−0.0003	0.0006	−0.0001	−0.0002	0.0002

Table 10: Coverage of nonzero parameter estimates ($R = 5$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.799	0.767	-	0.794	0.797	-	0.800	0.788	-	0.794	0.801
25	50	0.841	0.855	-	0.864	0.849	-	0.849	0.862	-	0.845	0.839
	100	0.885	0.879	-	0.875	0.887	0.879	0.877	0.892	-	0.871	0.871
	25	0.860	0.856	0.871	0.868	0.866	-	0.858	0.864	-	0.847	0.843
50	50	0.879	0.901	0.883	0.882	0.894	-	0.900	0.889	-	0.893	0.894
	100	0.912	0.909	0.886	0.915	0.902	0.907	0.913	0.905	-	0.900	0.904
	25	0.875	0.884	0.890	0.890	0.858	-	0.882	0.881	0.885	0.880	0.884
100	50	0.911	0.922	0.936	0.913	0.907	-	0.902	0.914	0.931	0.909	0.914
	100	0.927	0.919	0.933	0.941	0.926	0.940	0.937	0.940	0.926	0.896	0.903

Table 11: Bias of bias corrected estimates of nonzero parameters ($R = 10$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	−0.0001	−0.0020	-	−0.0009	0.0008	-	0.0020	0.0066	-	−0.0001	0.0006
25	50	−0.0004	−0.0014	-	0.0002	0.0006	-	0.0005	0.0052	-	−0.0004	0.0004
	100	−0.0001	−0.0002	-	0.0006	−0.0012	0.0014	−0.0010	0.0017	-	−0.0001	0.0001
	25	0.0003	−0.0005	−0.0001	0.0014	−0.0012	-	−0.0003	0.0015	-	−0.0004	0.0005
50	50	0	−0.0002	0	0.0003	−0.0010	-	−0.0008	0.0016	-	−0.0003	0.0002
	100	0	−0.0002	0	0.0003	−0.0006	0.0007	−0.0002	0.0006	-	−0.0002	0.0002
	25	−0.0001	−0.0004	0.0001	0.0005	−0.0010	-	0.0002	0.0030	−0.0011	−0.0003	0.0003
100	50	0.0002	−0.0004	0.0001	0.0005	−0.0009	-	−0.0008	0.0017	−0.0004	−0.0002	0.0002
	100	0.0001	−0.0002	0	0.0001	−0.0006	0.0005	−0.0003	0.0011	0	−0.0002	0.0002

Table 12: Coverage of nonzero parameter estimates ($R = 10$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.462	0.450	-	0.484	0.471	-	0.469	0.464	-	0.505	0.484
25	50	0.602	0.619	-	0.638	0.641	-	0.635	0.631	-	0.667	0.660
	100	0.739	0.728	-	0.721	0.720	0.738	0.714	0.724	-	0.711	0.703
	25	0.621	0.629	0.655	0.621	0.661	-	0.659	0.651	-	0.659	0.658
50	50	0.760	0.772	0.771	0.772	0.741	-	0.765	0.777	-	0.764	0.777
	100	0.823	0.829	0.822	0.839	0.818	0.828	0.833	0.810	-	0.840	0.843
	25	0.728	0.740	0.719	0.732	0.708	-	0.711	0.726	0.729	0.698	0.707
100	50	0.831	0.818	0.821	0.845	0.827	-	0.839	0.824	0.828	0.840	0.833
	100	0.853	0.863	0.887	0.869	0.871	0.884	0.868	0.877	0.870	0.851	0.851

Table 13: Percentage of true zeros ($R = 10$)

n	T	ρ_3	ρ_5	δ_2	δ_4	δ_{12}	δ_{14}	ϕ_2	ϕ_4	ϕ_5
	25	99.9	-	90	-	90.4	-	100	-	-
25	50	99.7	-	100	100	99.4	-	100	-	-
	100	99.7	-	100	100	100	-	99.7	-	-
	25	100	-	100	-	99.8	99.8	100	100	-
50	50	100	-	100	100	100	100	100	100	-
	100	99.6	-	100	100	100	100	99.6	99.6	-
	25	100	100	100	-	100	100	100	100	100
100	50	99.6	99.6	100	100	100	100	99.6	99.6	99.6
	100	99.6	99.6	99.9	99.9	99.9	99.9	99.7	99.6	99.6

Table 14: Percentage of false zeros ($R = 10$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.1	0.2	-	0	0	-	0	0	-	0	0
25	50	0.1	0.3	-	0	0	-	0	0	-	0	0
	100	0	0	-	0	0	0	0	0	-	0	0
	25	0	0	0	0	0	-	0	0	-	0	0
50	50	0	0	0	0	0	-	0	0	-	0	0
	100	0	0	0	0	0	0	0	0	-	0	0
	25	0	0	0	0	0	-	0	0	0	0	0
100	50	0	0	0	0	0	-	0	0	0	0	0
	100	0	0	0	0	0	0	0	0	0	0	0

J.2 Gamma(0.5,1)

Table 15: Bias of bias corrected estimates of nonzero parameters ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.0006	-0.0007	-	0.0009	-0.0006	-	-0.0027	0.0023	-	-0.0003	0.0004
25	50	0.0002	-0.0002	-	0.0006	-0.0002	-	-0.0026	0.0023	-	-0.0003	0.0004
	100	0.0001	-0.0003	-	0.0002	-0.0008	0.0008	0.0001	0.0009	-	-0.0001	0.0001
	25	0.0002	-0.0002	-0.0001	-0.0003	-0.0005	-	-0.0006	0.0015	-	-0.0003	0.0004
50	50	0.0002	-0.0004	0	0.0004	-0.0013	-	-0.0004	0.0012	-	-0.0002	0.0002
	100	0.0001	-0.0002	0.0001	-0.0001	-0.0002	0.0002	0	0.0003	-	-0.0002	0.0002
	25	0.0002	-0.0004	0.0002	0.0003	-0.0001	-	-0.0017	0.0028	-0.0014	-0.0003	0.0003
100	50	0.0001	-0.0002	0.0001	0.0004	-0.0003	-	-0.0007	0.0012	-0.0004	-0.0002	0.0002
	100	0.0001	-0.0001	0	0.0003	-0.0001	0.0001	-0.0005	0.0007	-0.0003	-0.0002	0.0002

Table 16: Coverage of nonzero parameter estimates ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.891	0.896	-	0.889	0.896	-	0.901	0.906	-	0.893	0.895
25	50	0.911	0.909	-	0.919	0.895	-	0.907	0.919	-	0.926	0.935
	100	0.912	0.922	-	0.922	0.916	0.914	0.920	0.920	-	0.929	0.926
	25	0.936	0.931	0.925	0.922	0.930	-	0.918	0.919	-	0.928	0.930
50	50	0.930	0.929	0.944	0.948	0.930	-	0.922	0.938	-	0.933	0.934
	100	0.930	0.925	0.940	0.940	0.929	0.926	0.945	0.943	-	0.952	0.952
	25	0.925	0.944	0.919	0.907	0.927	-	0.914	0.909	0.929	0.932	0.932
100	50	0.931	0.936	0.938	0.944	0.939	-	0.933	0.946	0.940	0.924	0.928
	100	0.944	0.936	0.942	0.947	0.951	0.946	0.935	0.948	0.929	0.935	0.936

Table 17: True number of factors is selected (%)

T	25			50			100		
n	IC1	IC2	IC3	IC1	IC2	IC3	IC1	IC2	IC3
25	9.3	69.4	33.3	26.4	71.8	62.5	63.4	75.3	71.4
50	33.2	67.9	59	30.7	73.7	65.1	60.6	77.9	73.4
100	58.3	70.7	66.5	56.2	75.2	69.2	51.9	80.9	73.5

J.3 Gamma(1,1)

Table 18: Bias of bias corrected estimates of nonzero parameters ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	−0.0001	−0.0004	-	0.0013	−0.0021	-	−0.0012	0.0022	-	−0.0001	0.0001
25	50	0.0001	−0.0003	-	0.0007	−0.0008	-	−0.0015	0.0015	-	−0.0003	0.0004
	100	0	0	-	−0.0001	−0.0001	0	0.0006	0.0010	-	−0.0001	0.0001
	25	0.0003	−0.0003	0.0001	0	−0.0001	-	−0.0018	0.0017	-	−0.0002	0.0002
50	50	0.0001	−0.0002	0.0001	0.0005	−0.0004	-	−0.0006	0.0008	-	−0.0003	0.0003
	100	0	−0.0002	0.0001	0.0001	−0.0005	0.0004	−0.0005	0.0011	-	−0.0001	0.0001
	25	0	0	0	0.0002	0.0003	-	−0.0005	0.0010	−0.0007	−0.0002	0.0002
100	50	0.0001	−0.0002	0.0001	0.0002	0	-	−0.0007	0.0011	−0.0006	−0.0002	0.0002
	100	0	0	0	0.0002	−0.0002	0.0001	−0.0003	0.0003	−0.0001	−0.0002	0.0002

Table 19: Coverage of nonzero parameter estimates ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.906	0.910	-	0.914	0.899	-	0.893	0.912	-	0.931	0.928
25	50	0.920	0.913	-	0.930	0.910	-	0.926	0.917	-	0.914	0.910
	100	0.912	0.918	-	0.925	0.934	0.936	0.914	0.925	-	0.932	0.936
	25	0.898	0.912	0.918	0.924	0.925	-	0.913	0.940	-	0.924	0.921
50	50	0.941	0.928	0.937	0.936	0.942	-	0.936	0.930	-	0.926	0.931
	100	0.934	0.947	0.938	0.942	0.935	0.949	0.942	0.943	-	0.943	0.944
	25	0.929	0.928	0.923	0.926	0.924	-	0.935	0.940	0.926	0.916	0.916
100	50	0.939	0.933	0.950	0.940	0.940	-	0.936	0.918	0.945	0.942	0.941
	100	0.936	0.934	0.932	0.944	0.943	0.953	0.945	0.938	0.938	0.939	0.939

Table 20: True number of factors is selected (%)

T	25			50			100		
n	IC1	IC2	IC3	IC1	IC2	IC3	IC1	IC2	IC3
25	7.5	79.6	50.2	32.8	80	72.8	71.9	82.7	80
50	37.4	76.4	69.1	36.9	84.6	75.6	61.1	82.2	77.2
100	61.3	75.2	70	63.1	85.3	79.8	52.5	90.1	80.7

J.4 Laplace(0,1)

Table 21: Bias of bias corrected estimates of nonzero parameters ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.0007	−0.0007	-	−0.0002	0.0001	-	−0.0039	0.0037	-	0	0
25	50	0.0003	−0.0005	-	0.0007	−0.0007	-	−0.0018	0.0024	-	−0.0003	0.0003
	100	0	−0.0001	-	0	−0.0007	0.0007	−0.0003	0.0009	-	−0.0002	0.0002
	25	0.0001	−0.0003	0	0.0002	−0.0007	-	−0.0004	0.0016	-	−0.0003	0.0003
50	50	0.0001	−0.0001	−0.0001	0.0004	−0.0003	-	−0.0012	0.0014	-	−0.0003	0.0003
	100	0	−0.0001	0	0.0003	−0.0006	0.0006	0.0004	0.0007	-	−0.0002	0.0002
	25	0	−0.0002	0	0.0005	−0.0012	-	0	0.0016	−0.0005	−0.0002	0.0002
100	50	0.0001	−0.0003	0.0001	0.0005	−0.0005	-	0.0010	0.0017	−0.0004	−0.0003	0.0003
	100	0	−0.0001	0	0.0002	−0.0001	0.0002	−0.0001	0.0006	−0.0005	−0.0002	0.0002

Table 22: Coverage of nonzero parameter estimates ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	0.900	0.914	-	0.905	0.890	-	0.913	0.903	-	0.906	0.896
25	50	0.913	0.905	-	0.928	0.921	-	0.925	0.926	-	0.919	0.917
	100	0.935	0.926	-	0.935	0.912	0.930	0.935	0.933	-	0.922	0.915
	25	0.917	0.932	0.926	0.921	0.929	-	0.908	0.923	-	0.909	0.907
50	50	0.934	0.926	0.939	0.934	0.940	-	0.928	0.918	-	0.942	0.938
	100	0.942	0.936	0.941	0.934	0.943	0.920	0.937	0.937	-	0.949	0.937
	25	0.939	0.927	0.911	0.931	0.934	-	0.934	0.925	0.929	0.927	0.926
100	50	0.935	0.926	0.935	0.931	0.925	-	0.933	0.938	0.942	0.924	0.927
	100	0.935	0.926	0.948	0.939	0.927	0.934	0.938	0.948	0.947	0.940	0.936

Table 23: True number of factors is selected (%)

T	25			50			100		
n	IC1	IC2	IC3	IC1	IC2	IC3	IC1	IC2	IC3
25	0	98.1	71.2	28.5	100	99.2	100	100	100
50	31.1	100	98.9	2.9	100	100	100	100	100
100	99.7	100	100	100	100	100	99.5	100	100

J.5 ChiSq(3)

Table 24: Bias of bias corrected estimates of nonzero parameters ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
	25	-0.0002	-0.0002	-	-0.0002	-0.0012	-	-0.0004	0.0025	-	-0.0003	0.0003
25	50	0.0001	-0.0003	-	0.0005	-0.0004	-	-0.0021	0.0023	-	-0.0003	0.0003
	100	0	-0.0001	-	0.0005	-0.0004	0.0005	-0.0009	0.0010	-	-0.0002	0.0002
	25	-0.0002	0	0	0.0006	-0.0008	-	0	0.0010	-	-0.0001	0.0001
50	50	-0.0001	0	0	0.0005	-0.0011	-	-0.0003	0.0010	-	-0.0003	0.0003
	100	0	-0.0001	0	0.0002	-0.0004	0.0004	-0.0006	0.0009	-	-0.0002	0.0002
	25	0	-0.0002	0.0001	0.0006	-0.0004	-	-0.0010	0.0015	-0.0006	-0.0002	0.0002
100	50	0	-0.0001	0.0001	0	-0.0005	-	-0.0010	0.0016	-0.0010	-0.0003	0.0003
	100	0	0	0	0.0002	-0.0002	0.0003	-0.0001	0.0002	-0.0003	-0.0003	0.0002

Table 25: Coverage of nonzero parameter estimates ($R = R^0$)

n	T	ρ_1	ρ_2	ρ_4	δ_1	δ_3	δ_5	δ_{11}	δ_{13}	δ_{15}	ϕ_1	ϕ_3
25	25	0.900	0.906	-	0.901	0.898	-	0.893	0.896	-	0.903	0.899
	50	0.908	0.923	-	0.901	0.927	-	0.927	0.924	-	0.917	0.912
	100	0.938	0.923	-	0.918	0.930	0.924	0.947	0.923	-	0.919	0.920
50	25	0.927	0.913	0.910	0.914	0.911	-	0.917	0.923	-	0.917	0.922
	50	0.922	0.936	0.932	0.925	0.921	-	0.920	0.934	-	0.931	0.929
	100	0.940	0.937	0.928	0.932	0.939	0.935	0.927	0.932	-	0.938	0.941
100	25	0.914	0.935	0.931	0.940	0.923	-	0.930	0.940	0.926	0.938	0.933
	50	0.942	0.935	0.931	0.935	0.931	-	0.946	0.922	0.952	0.932	0.932
	100	0.935	0.944	0.933	0.940	0.935	0.936	0.937	0.933	0.940	0.924	0.919

Table 26: True number of factors is selected (%)

T	25			50			100		
n	IC1	IC2	IC3	IC1	IC2	IC3	IC1	IC2	IC3
25	18.7	68.1	51.8	46.4	73.5	65.5	64.5	79.6	74.5
50	46.3	73	64.3	42.3	83.8	71.8	61.7	78.8	73.1
100	61.7	78.8	73.1	63.2	84.6	79.5	55.5	92.5	85.7

K Additional Simulations

As has been touched on several times throughout the paper, since both common factors and interaction generate dependence in the cross-section, it can sometime be difficult to disentangle these two effects. Proposition ID provides an argument to demonstrate that, asymptotically at least, it is possible to separate out these two features, under certain conditions. However, for the interested reader, this section provides results for two additional Monte Carlo experiments which are designed to assess how well the method might perform in settings where separating the effects of interaction and of common factors may be especially difficult.

K.1 Pure Star

In this design there are no exogenous covariates and only a single (row normalised) weights matrix with associated coefficient $\rho^0 = 0.2$. The network consists of a star, where all the cross-sectional units are connected to the first unit and to no others. This produces a corresponding weights matrix which will always have a rank of 2. As in the main text, the factors, loadings and errors are generated as standard normal, however the true factor term has a rank of 2, that is $R^0 = 2$. With the factor term and the weights matrix both having a low rank, and with exogenous covariates absent, this provides an especially challenging design. The following tables provide results with the postulated number of factors R being correctly specified, and overspecified to various degrees.

Table 27: Bias of bias corrected estimates

	R			
$n = T$	2	3	4	6
25	-0.0072	-0.0331	-0.0616	-0.1516
50	-0.0001	-0.0100	-0.0200	-0.0516
100	-0.0001	-0.0006	-0.0022	-0.0080

Table 28: Coverage

	R			
$n = T$	2	3	4	6
25	0.844	0.676	0.529	0.222
50	0.889	0.796	0.722	0.578
100	0.891	0.899	0.834	0.737

Table 29: Percentage of false zeros

	R			
$n = T$	2	3	4	6
25	5.2	13.8	19.7	21.3
50	0.1	4.6	10.2	20.7
100	0	0.2	1	3.7

In this experiment, the number of factors being overspecified has a substantial influence on the performance of the procedure. Most telling are perhaps the results presented in Table 29 which give the percentage of times, across the Monte Carlo draws, that the coefficient ρ is incorrectly set to zero. With a small sample, there is an especially large increase in the number of false zeros once R exceeds 4. This might be explained by 4 being the combined rank of both the true factor term and the weights matrix. In all cases, however, the percentage of false zeros dramatically deceases as sample size increases.

K.2 Multiple Stars

This experiment is designed to more closely resemble the Monte Carlo design in the main text, with the number of weights matrices increasing with sample size. It is summarised in Table 30.

Table 30: True parameter values

$n=T$	ρ_1^0	ρ_2^0	ρ_3^0	ρ_4^0	ρ_5^0	δ_1^0	δ_2^0	δ_3^0	ϕ_1^0	ϕ_2^0	ϕ_3^0	ϕ_4^0	ϕ_5^0
25	0.2	0.2	0	-	-	3	0	-3	0.15	0	-0.15	$-$	$-$
50	0.2	0.2	0	0.2	-	3	0	-3	0.15	0	-0.15	0	$-$
100	0.2	0.2	0	0.2	0	3	0	-3	0.15	0	-0.15	0	0

Exogenous covariates are included in the model with these being generated according to $x_{\kappa it}^* = \nu + \sum_{r=1}^{R^0} \lambda_{ir}^0 f_{rt}^0 + e_{it}$ with ν being uniformly drawn from the integers $\{-10, ..., 10\}$ and $e_{it} \sim \mathcal{N}(0, 2)$, as in the main text. The factors, loadings and errors are all standard normal. The weights matrices take the form of stars, as in the previous experiment, however, these stars are of sizes $5, 7, 13, 25$ and 50. Tables 31–34 below summarise the results with the true number of factors (R^0) equal to 3 and the number of factors used in estimation (R) is correctly specified, or overspecified to varying degrees as $5, 10$ and 15.

Table 31: Bias of bias corrected estimates of nonzero parameters

R	$n=T$	ρ_1^0	ρ_2^0	ρ_4^0	δ_1^0	δ_3^0	ϕ_1^0	ϕ_3^0
	25	-0.0006	-0.0008	-	0.0002	-0.0001	-0.0001	0.0005
3	50	-0.0002	-0.0001	0	0.0001	-0.0002	0	0.0001
	100	0	-0.0001	0	-0.0001	0.0001	0	0.0002
	25	-0.0008	-0.0006	-	-0.0005	-0.0008	-0.0002	0.0005
5	50	-0.0001	0	-0.0001	0.0004	-0.0001	0	0.0001
	100	-0.0001	-0.0001	0	0.0001	0	0	0.0002
	25	-0.0107	-0.0110	-	-0.0024	0.0012	-0.0004	0.0050
10	50	-0.0002	-0.0005	-0.0004	0.0005	-0.0002	0	0.0009
	100	0	0	0	0	0	0	0.0002
	25	-0.0287	-0.0297	-	-0.0033	0.0016	-0.0010	0.0144
15	50	-0.0017	-0.0025	-0.0008	0.0003	-0.0003	0	0.0029
	100	-0.0001	-0.0006	-0.0001	0	0	0	0.0007

Table 32: Coverage of nonzero parameter estimates

R	$n = T$	ρ_1^0	ρ_2^0	ρ_4^0	δ_1^0	δ_3^0	ϕ_1^0	ϕ_3^0
3	25	0.805	0.795	-	0.804	0.806	0.796	0.798
	50	0.916	0.897	0.896	0.892	0.892	0.895	0.888
	100	0.942	0.928	0.930	0.926	0.928	0.917	0.910
5	25	0.789	0.808	-	0.796	0.744	0.807	0.788
	50	0.890	0.889	0.888	0.894	0.891	0.881	0.889
	100	0.918	0.929	0.935	0.917	0.909	0.920	0.914
10	25	0.513	0.454	-	0.436	0.481	0.457	0.504
	50	0.794	0.770	0.790	0.783	0.786	0.759	0.777
	100	0.874	0.893	0.889	0.891	0.883	0.873	0.870
15	25	0.225	0.237	-	0.211	0.217	0.231	0.207
	50	0.610	0.581	0.612	0.620	0.602	0.623	0.602
	100	0.786	0.824	0.813	0.836	0.822	0.825	0.829

Table 33: Percentage of false zeros

R	$n = T$	ρ_3^0	ρ_5^0	δ_3^0	ϕ_2^0	ϕ_4^0	ϕ_5^0
3	25	100	–	100	100	–	–
	50	100	–	100	100	100	–
	100	100	100	100	100	100	100
5	25	100	–	100	100	–	–
	50	100	–	100	100	100	–
	100	100	100	100	100	100	100
10	25	99.2	–	98.2	97.7	–	–
	50	100	–	100	100	100	–
	100	100	100	100	100	100	100
15	25	86.2	–	77.3	84.5	–	–
	50	100	–	100	100	100	–
	100	100	100	100	100	100	100

Table 34: Percentage of true zeros

R	$n = T$	ρ_1^0	ρ_2^0	ρ_4^0	δ_1^0	δ_3^0	ϕ_1^0	ϕ_3^0
3	25	0.2	0.2	-	0	0	0	0.2
	50	0	0	0	0	0	0	0
	100	0	0	0	0	0	0	0
5	25	0.1	0.1	-	0	0	0	0.2
	50	0	0	0	0	0	0	0
	100	0	0	0	0	0	0	0
10	25	4.7	4.6	-	0	0	0	5.5
	50	0.1	0.1	0	0	0	0	0.5
	100	0	0	0	0	0	0	0
15	25	12.3	10.5	-	0	0	0	11.2
	50	0.7	0.7	0.3	0	0	0	1.9
	100	0	0.2	0	0	0	0	0.3

A similar pattern emerges in this experiment to that in the previous section; overspecification of the number of factors can have a substantial effect on the significance of coefficients in small samples. In this case the impact is less poignant, however the same result is bourne out: as sample size increases, the ability of the procedure to separate the network structure and the factor term rapidly improves.

References

Bai, J., 2009. Panel data models with interactive fixed effects. Econometrica 77 (4), 1229–1279.

Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. Journal of Econometrics 150 (1), 41–55.

Davidson, J., 1994. Stochastic Limit Theory. Oxford University Press, Oxford, England.

Fan, J., Peng, H., 2004. Nonconcave penalized likelihood with a diverging number of parameters. The Annals of Statistics 32 (3), 928–961.

Horn, R. A., Johnson, C. R., 2012. Matrix Analysis, 2nd Edition. Cambridge University Press, New York, USA.

Kwok, H. H., 2019. Identification and estimation of linear social interaction models. Journal of Econometrics 210 (2), 434–458.

Latala, R., 2005. Some estimates of norms of random matrices. Proceedings of the American Mathematical Society 133 (5), 1273–1282.

Lee, L.-F., Yu, J., 2016. Identification of spatial durbin panel models. Journal of Applied Econometrics 31 (1), 133–162.

Moon, H. R., Weidner, M., 2015. Linear regression for panel with unknown number of factors as interactive fixed effects. Econometrica 83 (4), 1543–1579.

Moon, H. R., Weidner, M., 2017. Dynamic linear panel regression models with interactive fixed effects. Econometric Theory 33 (1), 158–195.

Shi, W., Lee, L.-F., 2017. Spatial dynamic panel data models with interactive fixed effects. Journal of Econometrics 197 (2), 323–347.

Yu, J., de Jong, R., Lee, L.-F., 2008. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. Journal of Econometrics 146 (1), 118–134.