

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

Ayden Higgins
University of Surrey, UK

Federico Martellosio
University of Surrey, UK

November 22, 2021

Contents

B Identification	1
C Assumption 4.2 and Equations (13) and (14)	3
D Proofs of Propositions 1 and 4	4
E Proofs of Lemmas A.1–A.3	13
F Proofs of Lemmas A.4–A.6	18
G Proofs of Lemmas F.1–F.4	44
H Proof of Proposition ID	102
I Verifying Assumptions 1–8	105
J Additional Tables	106
K Additional Simulations	113

B Identification

Further to the discussion at the start of Section 3, a few additional remarks are provided regarding identification. The term “identification” is used to refer to the condition that the distribution of the data is unique at the true parameter values. A related identification concept usually encountered when studying extremum estimators requires that the limit of the objective function is uniquely maximised at the true parameter values. This may be

referred to as *extremum based* identification (Lewbel, 2019). In combination with uniform convergence of the sample objective function, extremum based identification is used to establish the consistency of an extremum estimator in an archetypical manner; e.g., Newey and McFadden (1994), Theorem 2.1. The basic intuition behind this approach is that “*the limit of the maximum $\hat{\theta}$ should be the maximum of the limit*”, with the latter being unique (Newey and McFadden, 1994, p. 2120). In models with an increasing number of parameters this intuition becomes difficult to apply because the distribution of the data may change with each observation, and therefore there is typically no fixed population objective function towards which a sequence of sample objective functions can be said to converge. Nonetheless, the loss of intuition does neither preclude consistency nor identification. In the present case consistency is established in Proposition 1, while Assumption ID below sets out conditions that, in addition to Assumptions 1 and 2, can be used to establish identification.

Assumption ID.

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

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

ID.3 $\mu_1(\mathbb{E}[\mathbf{Z}'\mathbf{Z}]) \leq c < \infty$.

Notice that Assumption ID.1 is the same as Assumption 4.1, and Assumptions ID.2 and ID.3 are counterparts to Assumptions 4.2 and 4.3, and thus share analogous intuition.

Proposition ID (Identification). *Under Assumptions 1,2 and ID, the parameters θ^0, σ_0^2 and the product $\Lambda^0 \mathbf{F}^{0'}$ are identified.*

The proof of Proposition ID follows a similar construction to Theorem 2.1 in Moon and Weidner (2015) and establishes that for each (n, T) such that Assumption ID is satisfied, the expected quasi-likelihood is uniquely maximised at the true parameters.¹ This in turn implies that for each of these (n, T) the distribution of the data is unique at $(\theta^0, \sigma_0^2, \Lambda^0 \mathbf{F}^{0'})$ and thereby establishes identification. Moreover, the true number of factors R^0 can be recovered from the rank of the matrix $\Lambda^0 \mathbf{F}^{0'}$. One may ask if, for the present model, it is possible to follow the Bramoullé et al. (2009) approach and derive identification conditions in terms of the network structure (see also Kwok, 2019, for the case of multiple weights matrices). 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).

¹Since $\text{rank}(\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda}) \leq (T - R^0)(n - R)$ this obviously requires $(T - R^0)(n - R) \geq P$.

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

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

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

$$\begin{aligned} &\inf_{\mathbf{\Lambda} \in \mathbb{R}^{n \times R}, \mathbf{F} \in \mathbb{R}^{T \times R^0}} \mu_P(\mathcal{H}_1(\mathbf{\Lambda}, \mathbf{F})) \\ &= \min_{\boldsymbol{\alpha} \in \mathbb{R}^P: \|\boldsymbol{\alpha}\|_2=1} \inf_{\mathbf{\Lambda} \in \mathbb{R}^{n \times R}, \mathbf{F} \in \mathbb{R}^{T \times R^0}} \frac{1}{nT} (\mathbf{Z}\boldsymbol{\alpha})' (\mathbf{M}_{\mathbf{F}} \otimes \mathbf{M}_{\mathbf{\Lambda}}) \mathbf{Z}\boldsymbol{\alpha} \\ &= \min_{\boldsymbol{\alpha} \in \mathbb{R}^P: \|\boldsymbol{\alpha}\|_2=1} \inf_{\mathbf{\Lambda} \in \mathbb{R}^{n \times R}, \mathbf{F} \in \mathbb{R}^{T \times R^0}} \frac{1}{nT} \text{tr} \left((\boldsymbol{\alpha} \cdot \mathbf{Z})' \mathbf{M}_{\mathbf{\Lambda}} (\boldsymbol{\alpha} \cdot \mathbf{Z}) \mathbf{M}_{\mathbf{F}} \right) \\ &= \min_{\boldsymbol{\alpha} \in \mathbb{R}^P: \|\boldsymbol{\alpha}\|_2=1} \sum_{r=R+R^0+1}^n \mu_r \left(\frac{1}{nT} (\boldsymbol{\alpha} \cdot \mathbf{Z})(\boldsymbol{\alpha} \cdot \mathbf{Z})' \right), \end{aligned} \quad (\text{C.2})$$

where the last line follows from Lemma A.1 in Moon and Weidner (2017). The inequality in (C.1) follows since $\inf_{\mathbf{\Lambda} \in \mathbb{R}^{n \times R}} \mu_P(\mathcal{H}_1(\mathbf{\Lambda}, \mathbf{F}^0))$ 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 \mathbf{F}^0 . Assumption A in Bai (2009) is equivalent to the requirement that $\inf_{\mathbf{\Lambda} \in \mathbb{R}^{n \times R}} \mu_P(\mathcal{H}_1(\mathbf{\Lambda}, \mathbf{F}^0)) > 0$. 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 behind these is that Assumptions 4.2 and 4.3 provide lower and upper bounds on variation in the data, respectively, from which the inequalities (13) and (14) can be derived. It would usually be assumed that the matrix $\frac{1}{nT} \mathbf{Z}' \mathbf{Z}$ is positive definite in the limit. However, it is shown next that the p -th eigenvalue of $\frac{1}{nT} \mathbf{Z}' \mathbf{Z}$ can be no less than the p -th eigenvalue of $\frac{1}{nT} \mathbf{Z}' (\mathbf{M}_{\mathbf{F}} \otimes \mathbf{M}_{\mathbf{\Lambda}}) \mathbf{Z}$ whereby

Assumption 4.2 implies (13) and Assumption 4.3 implies (14). Let $\mathbf{M} := \mathbf{M}_F \otimes \mathbf{M}_\Lambda$. Since the Kronecker product of two symmetric and idempotent matrices is also symmetric and idempotent, both \mathbf{M} and $\mathbf{P} := \mathbf{I}_{n \times T} - \mathbf{M}$ are symmetric and idempotent.² From Weyl's inequality, for two $n \times n$ symmetric matrices \mathbf{A}, \mathbf{B} of the same size $\mu_i(\mathbf{A}) + \mu_n(\mathbf{B}) \leq \mu_i(\mathbf{A} + \mathbf{B})$ for $i = 1, \dots, n$ (e.g., Horn and Johnson, 2012, Corollary 4.3.15). As $\mathbf{Z}'\mathbf{Z} = \mathbf{Z}'\mathbf{M}\mathbf{Z} + \mathbf{Z}'\mathbf{P}\mathbf{Z}$, and all three of these matrices are real and symmetric, then for $p = 1, \dots, P$,

$$\mu_p \left(\frac{1}{nT} \mathbf{Z}'\mathbf{M}\mathbf{Z} \right) + \mu_P \left(\frac{1}{nT} \mathbf{Z}'\mathbf{P}\mathbf{Z} \right) \leq \mu_p \left(\frac{1}{nT} \mathbf{Z}'\mathbf{Z} \right). \quad (\text{C.3})$$

By Assumption 4.2, $\mu_P \left(\frac{1}{nT} \mathbf{Z}'\mathbf{M}\mathbf{Z} \right) \geq c_1 > 0$ w.p.a.1. Also $\mu_P \left(\frac{1}{nT} \mathbf{Z}'\mathbf{P}\mathbf{Z} \right) \geq 0$ since \mathbf{P} is idempotent, and therefore $\frac{1}{nT} \mathbf{Z}'\mathbf{P}\mathbf{Z}$ must be positive semidefinite. Hence (14) follows from (C.3). Similarly, (13) follows from (C.3) since, by Assumption 4.3, $\mu_1 \left(\frac{1}{nT} \mathbf{Z}'\mathbf{Z} \right) \leq c_2 < \infty$ w.p.a.1.

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.

Consistency of the QMLE $\tilde{\boldsymbol{\theta}}$

The average concentrated quasi-likelihood is given by

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

The first step involves establishing a lower bound for $\mathcal{L}(\boldsymbol{\theta}^0)$. Evaluating (D.1) at $\boldsymbol{\theta}^0$ and substituting in the true data generating process yields

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

Now,

$$0 \leq \inf_{\boldsymbol{\Lambda} \in \mathbb{R}^{n \times R}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \boldsymbol{\varepsilon}_t)' \mathbf{M}_\Lambda (\boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \boldsymbol{\varepsilon}_t) \right\}$$

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

$$\begin{aligned}
&\leq \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \boldsymbol{\varepsilon}_t)' \mathbf{M}_{\boldsymbol{\Lambda}^0} (\boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \boldsymbol{\varepsilon}_t) \\
&= \frac{1}{nT} \sum_{t=1}^T \boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t - \frac{1}{nT} \sum_{t=1}^T \boldsymbol{\varepsilon}_t' \mathbf{P}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon}_t.
\end{aligned} \tag{D.3}$$

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.3),

$$\left| \frac{1}{nT} \sum_{t=1}^T \boldsymbol{\varepsilon}_t' \mathbf{P}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon}_t \right| = \frac{1}{nT} |\text{tr}(\boldsymbol{\varepsilon}' \mathbf{P}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon})| \leq \frac{1}{nT} R^0 \|\boldsymbol{\varepsilon}\|_2^2 = O_P\left(\frac{1}{\min\{n, T\}}\right). \tag{D.4}$$

This gives the result that

$$\begin{aligned}
\underline{\mathcal{L}}(\boldsymbol{\theta}^0) &:= \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log\left(\sigma_0^2 + O_P\left(\frac{1}{\min\{n, T\}}\right)\right) \\
&= \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2 + O_P(a_{nT}^2)) \leq \mathcal{L}(\boldsymbol{\theta}^0).
\end{aligned} \tag{D.5}$$

Now, in the second step an upper bound for $\mathcal{L}(\boldsymbol{\theta})$ is established. Substituting the true data generating process into (D.1) yields

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\theta}) &= \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) - \frac{1}{2} \log\left(\inf_{\boldsymbol{\Lambda} \in \mathbb{R}^{n \times R}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} (\mathbf{X}_t \boldsymbol{\beta}^0 + \boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \boldsymbol{\varepsilon}_t) - \mathbf{X}_t \boldsymbol{\beta})' \right. \right. \\
&\quad \left. \left. \times \mathbf{M}_{\boldsymbol{\Lambda}} (\mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} (\mathbf{X}_t \boldsymbol{\beta}^0 + \boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \boldsymbol{\varepsilon}_t) - \mathbf{X}_t \boldsymbol{\beta}) \right\}\right) \\
&\leq \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) - \frac{1}{2} \log\left(\inf_{\dot{\boldsymbol{\Lambda}} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \mathbf{X}_t \boldsymbol{\beta}^0 + \mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\varepsilon}_t - \mathbf{X}_t \boldsymbol{\beta})' \right. \right. \\
&\quad \left. \left. \times \mathbf{M}_{\dot{\boldsymbol{\Lambda}}} (\mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \mathbf{X}_t \boldsymbol{\beta}^0 + \mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\varepsilon}_t - \mathbf{X}_t \boldsymbol{\beta}) \right\}\right),
\end{aligned} \tag{D.6}$$

where the last expression is obtained by also minimising with respect to $\mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\Lambda}^0$ and \mathbf{F}^0 since the value of the objective function in curly brackets can be no less than if one was also able to minimise over $\mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\Lambda}^0$ and \mathbf{F}^0 (e.g, Moon and Weidner, 2015). Lemma A.1 in Moon and Weidner (2017) then demonstrates the equivalence between this and the second expression as it appears in (D.6), where the expression is now minimised over $\dot{\boldsymbol{\Lambda}} \in \mathbb{R}^{n \times (R+R^0)}$ because the rank of $\mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\Lambda}^0 \mathbf{F}^{0'} - \boldsymbol{\Lambda} \mathbf{F}'$ can be no greater than $R + R^0$. Applying Lemma A.2(i) to (D.6) gives

$$\mathcal{L}(\boldsymbol{\theta}) \leq \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) - \frac{1}{2} \log\left(\inf_{\dot{\boldsymbol{\Lambda}} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) + \mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\varepsilon}_t)' \right. \right.$$

$$\times M_{\hat{\Lambda}}(\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) + \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t) \Bigg\} \Bigg). \quad (\text{D.7})$$

Expanding the term inside of the log in (D.7),

$$\begin{aligned} & \inf_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) + \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t)' M_{\hat{\Lambda}}(\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) + \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t) \right\} \\ & \geq \inf_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}))' M_{\hat{\Lambda}} \mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) \right\} + \frac{2}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}))' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t \\ & \quad - \sup_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{2}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}))' \mathbf{P}_{\hat{\Lambda}} \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t \right\} + \frac{1}{nT} \sum_{t=1}^T (\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t)' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t \\ & \quad - \sup_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t)' \mathbf{P}_{\hat{\Lambda}} \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t \right\} \\ & =: k_1 + \dots + k_5. \end{aligned} \quad (\text{D.8})$$

Consider the probability order of terms k_1, \dots, k_5 .

$$\begin{aligned} k_1 &= \inf_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{\theta}^0 - \boldsymbol{\theta})' \mathbf{Z}_t' M_{\hat{\Lambda}} \mathbf{Z}_t (\boldsymbol{\theta}^0 - \boldsymbol{\theta}) \right\} \\ &\geq \inf_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \mu_P \left(\frac{1}{nT} \boldsymbol{\mathcal{Z}}' (\mathbf{I}_T \otimes M_{\hat{\Lambda}}) \boldsymbol{\mathcal{Z}} \right) \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2^2 \geq c_1 \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2^2 > 0, \end{aligned} \quad (\text{D.9})$$

where the last inequality holds with probability approaching 1 as $n, T \rightarrow \infty$ by Assumption 4.2. To see this, notice

$$\begin{aligned} & \inf_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \mu_P \left(\frac{1}{nT} \boldsymbol{\mathcal{Z}}' (\mathbf{I}_T \otimes M_{\hat{\Lambda}}) \boldsymbol{\mathcal{Z}} \right) \\ &= \min_{\boldsymbol{\alpha} \in \mathbb{R}^P: \|\boldsymbol{\alpha}\|_2=1} \inf_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \frac{1}{nT} (\boldsymbol{\mathcal{Z}}\boldsymbol{\alpha})' (\mathbf{I}_T \otimes M_{\hat{\Lambda}}) \boldsymbol{\mathcal{Z}}\boldsymbol{\alpha} \\ &= \min_{\boldsymbol{\alpha} \in \mathbb{R}^P: \|\boldsymbol{\alpha}\|_2=1} \inf_{\hat{\Lambda} \in \mathbb{R}^{n \times (R+R^0)}} \frac{1}{nT} \text{tr} \left((\boldsymbol{\alpha} \cdot \boldsymbol{\mathcal{Z}})' M_{\hat{\Lambda}} (\boldsymbol{\alpha} \cdot \boldsymbol{\mathcal{Z}}) \right) \\ &= \min_{\boldsymbol{\alpha} \in \mathbb{R}^P: \|\boldsymbol{\alpha}\|_2=1} \sum_{r=R+R^0+1}^n \mu_r \left(\frac{1}{nT} (\boldsymbol{\alpha} \cdot \boldsymbol{\mathcal{Z}}) (\boldsymbol{\alpha} \cdot \boldsymbol{\mathcal{Z}})' \right), \end{aligned} \quad (\text{D.10})$$

where (D.10) is shown to be equivalent to $\inf_{\boldsymbol{\Lambda} \in \mathbb{R}^{n \times R}, \mathbf{F} \in \mathbb{R}^{T \times R^0}} \mu_P(\mathcal{H}_1(\boldsymbol{\Lambda}, \mathbf{F}))$ in (C.2) which in turn is bounded away from zero with probability approaching 1 by Assumption 4.2. Next,

$$|k_2| = \frac{2}{nT} \sum_{p=1}^P |\theta_p^0 - \theta_p| |\text{tr}(\boldsymbol{\mathcal{Z}}_p' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon})| \leq \frac{2}{nT} \left(\sum_{p=1}^P |\theta_p^0 - \theta_p|^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^P |\text{tr}(\boldsymbol{\mathcal{Z}}_p' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon})|^2 \right)^{\frac{1}{2}}$$

$$= \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2 O_P \left(\sqrt{\frac{P}{nT}} \right), \quad (\text{D.11})$$

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

$$\begin{aligned} |k_3| &= \left| \sup_{\dot{\boldsymbol{\Lambda}} \in \mathbb{R}^{n \times (R+R^0)}} \left\{ \frac{2}{nT} \sum_{p=1}^P (\theta_p^0 - \theta_p) \text{tr}(\boldsymbol{z}_p' \boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}} \boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}) \right\} \right| \\ &\leq \frac{2(R+R^0)}{nT} \sum_{p=1}^P \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2 \|\boldsymbol{z}_p\|_2 \|\boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1}\|_2 \|\boldsymbol{\varepsilon}\|_2 \\ &\leq \frac{2(R+R^0)}{nT} \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2 \|\boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1}\|_2 \|\boldsymbol{\varepsilon}\|_2 \left(\sum_{p=1}^P \|\boldsymbol{z}_p\|_2^2 \right)^{\frac{1}{2}} \\ &= \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2 O_P \left(\sqrt{\frac{P}{\min\{n, T\}}} \right), \end{aligned} \quad (\text{D.12})$$

which follows because $\|\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}}\|_2 = 1$, since the maximum eigenvalue of any projection matrix is 1, $\|\boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1}\|_2 \leq \sqrt{\|\boldsymbol{S}(\boldsymbol{\rho})\|_1 \|\boldsymbol{S}(\boldsymbol{\rho})\|_\infty} \sqrt{\|\boldsymbol{S}^{-1}\|_1 \|\boldsymbol{S}^{-1}\|_\infty}$ and both $\boldsymbol{S}(\boldsymbol{\rho})$ 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,

$$k_4 = \frac{\sigma_0^2}{n} \text{tr}((\boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1})' \boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1}) + O_P \left(\frac{1}{\sqrt{nT}} \right) \quad (\text{D.13})$$

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

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

and thus

$$\begin{aligned} |k_5| &\leq \frac{(R+R^0)}{nT} \|\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}} \boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1} \boldsymbol{\varepsilon}\|_2^2 \\ &\leq \frac{(R+R^0)}{nT} \|\boldsymbol{S}(\boldsymbol{\rho}) \boldsymbol{S}^{-1}\|_2^2 \|\boldsymbol{\varepsilon}\|_2^2 = O_P \left(\frac{1}{\min\{n, T\}} \right), \end{aligned} \quad (\text{D.14})$$

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

$$\mathcal{L}(\boldsymbol{\theta}) \leq \frac{1}{n} \log(\det(\boldsymbol{S}(\boldsymbol{\rho}))) - \frac{1}{2} \log \left(c_1 \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2^2 + O_P \left(\frac{1}{\min\{n, T\}} \right) \right)$$

$$\begin{aligned}
& + \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) + O_P\left(\frac{1}{\sqrt{nT}}\right) + \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2 O_P\left(\sqrt{\frac{P}{nT}}\right)) \\
& = \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) - \frac{1}{2} \log\left(c_1 \|\boldsymbol{\theta} - \boldsymbol{\theta}^0\|_2^2 + O_P(a_{nT}) \|\boldsymbol{\theta} - \boldsymbol{\theta}^0\|_2 + O_P(a_{nT}^2)\right) \\
& \quad + \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) \\
& =: \bar{\mathcal{L}}(\boldsymbol{\theta}). \tag{D.15}
\end{aligned}$$

Now, since $\tilde{\boldsymbol{\theta}}$ is a global maximiser, $\mathcal{L}(\boldsymbol{\theta}^0) \leq \mathcal{L}(\tilde{\boldsymbol{\theta}})$ and therefore $\underline{\mathcal{L}}(\boldsymbol{\theta}^0) \leq \bar{\mathcal{L}}(\tilde{\boldsymbol{\theta}})$. Using the expressions for these bounds derived in (D.5) and (D.15) gives

$$\begin{aligned}
& \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2 + O_P(a_{nT}^2)) \\
& \leq \frac{1}{n} \log(\det(\mathbf{S}(\tilde{\boldsymbol{\rho}}))) - \frac{1}{2} \log\left(c_1 \|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2^2 + O_P(a_{nT}) \|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 + O_P(a_{nT}^2)\right) \\
& \quad + \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\tilde{\boldsymbol{\rho}})\mathbf{S}^{-1})'\mathbf{S}(\tilde{\boldsymbol{\rho}})\mathbf{S}^{-1}). \tag{D.16}
\end{aligned}$$

Multiplying both sides of (D.16) by -2 , exponentiating, and then noticing that, by Lemma A.1, $\sigma_0^2 \det((\mathbf{S}(\tilde{\boldsymbol{\rho}})\mathbf{S}^{-1})'\mathbf{S}(\tilde{\boldsymbol{\rho}})\mathbf{S}^{-1})^{\frac{1}{n}} \leq \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\tilde{\boldsymbol{\rho}})\mathbf{S}^{-1})'\mathbf{S}(\tilde{\boldsymbol{\rho}})\mathbf{S}^{-1})$, results in

$$0 \geq c_1 \|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2^2 + O_P(a_{nT}) \|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 + O_P(a_{nT}^2). \tag{D.17}$$

Completing the square, $0 \geq (\sqrt{c_1} \|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 + O_P(a_{nT}))^2 + O_P(a_{nT}^2)$, whereby it follows that $\|\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 = O_P(a_{nT})$.

Consistency of the PQMLE $\hat{\boldsymbol{\theta}}$

Since $\hat{\boldsymbol{\theta}}$ is the maximiser of the penalised quasi-likelihood function, $\mathcal{Q}(\boldsymbol{\theta}^0) \leq \mathcal{Q}(\hat{\boldsymbol{\theta}})$. Thus,

$$\begin{aligned}
\mathcal{Q}(\boldsymbol{\theta}^0) & = \mathcal{L}(\boldsymbol{\theta}^0) - \left(\gamma_\rho \sum_{q=1}^Q \omega_q |\rho_q^0| + \gamma_\beta \sum_{k=1}^K \omega_{Q+k} |\beta_k^0| \right) \\
& \leq \mathcal{Q}(\hat{\boldsymbol{\theta}}) \\
& = \mathcal{L}(\hat{\boldsymbol{\theta}}) - \left(\gamma_\rho \sum_{q=1}^Q \omega_q |\hat{\rho}_q| + \gamma_\beta \sum_{k=1}^K \omega_{Q+k} |\hat{\beta}_k| \right) \\
& \leq \mathcal{L}(\hat{\boldsymbol{\theta}}). \tag{D.18}
\end{aligned}$$

Consider the penalty term. Using Assumption 3.1,

$$\gamma_\rho \sum_{q=1}^Q \omega_q |\rho_q^0| + \gamma_\beta \sum_{k=1}^K \omega_{Q+k} |\beta_k^0| \leq \max\{\gamma_\rho, \gamma_\beta\} \sum_{p=1}^P \omega_p |\theta_p^0|$$

$$\begin{aligned}
&= \max\{\gamma_\rho, \gamma_\beta\} \sum_{p=1}^P \frac{1}{|\theta_p^\dagger|^\zeta} |\theta_p^0| \\
&\leq c_2 \max\{\gamma_\rho, \gamma_\beta\} P^0 \frac{1}{|\theta_{\underline{p}}|^\zeta} \\
&= c_2 \max\{\gamma_\rho, \gamma_\beta\} P^0 \left(\left| \frac{\theta_{\underline{p}}^\dagger}{\theta_{\underline{p}}^0} \right| \right)^{-\zeta} |\theta_{\underline{p}}^0|^{-\zeta}, \tag{D.19}
\end{aligned}$$

where $\underline{p} := \arg \min_{1 \leq p \leq P: \theta_p^0 \neq 0} |\theta_p^\dagger|$. Since the initial estimate $\boldsymbol{\theta}^\dagger$ satisfies $\|\boldsymbol{\theta}^\dagger - \boldsymbol{\theta}^0\|_2 = O_P(r_{nT}) = o_P(1)$, it follows that $|\theta_{\underline{p}}^\dagger/\theta_{\underline{p}}^0 - 1| \leq \frac{1}{|\theta_{\underline{p}}^0|} \|\boldsymbol{\theta}^\dagger - \boldsymbol{\theta}^0\|_2 = o_P(1)$ which implies $\theta_{\underline{p}}^\dagger/\theta_{\underline{p}}^0 = O_P(1)$. Hence,

$$\gamma_\rho \sum_{q=1}^Q \omega_q |\rho_q^0| + \gamma_\beta \sum_{k=1}^K \omega_{Q+k} |\beta_k^0| = \max\{\gamma_\rho, \gamma_\beta\} O_P(P^0) = O_P(a_{nT}^2), \tag{D.20}$$

under Assumption 3. Next, using (D.20), and applying the lower and upper bounds derived in (D.5) and (D.15) to (D.18) gives

$$\begin{aligned}
&\frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2 + O_P(a_{nT}^2)) + O_P(a_{nT}^2) \leq \frac{1}{n} \log(\det(\mathbf{S}(\hat{\boldsymbol{\rho}}))) \\
&- \frac{1}{2} \log \left(c_1 \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2^2 + O_P(a_{nT}) \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 + O_P(a_{nT}^2) + \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})' \mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1}) \right). \tag{D.21}
\end{aligned}$$

After rearranging and simplifying this becomes

$$\begin{aligned}
&\log \left(\sigma_0^2 \det((\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})' \mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})^{\frac{1}{n}} + O_P(a_{nT}^2) \right) + O_P(a_{nT}^2) \\
&\geq \log \left(c_1 \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2^2 + O_P(a_{nT}) \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 + O_P(a_{nT}^2) + \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})' \mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1}) \right). \tag{D.22}
\end{aligned}$$

Exponentiating, using the fact that by Assumption 4.4 $O_P(a_{nT}^2) = o_P(1)$, and Lemma A.1 gives the result

$$0 \geq c_1 \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2^2 + O_P(a_{nT}) \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 + O_P(a_{nT}^2), \tag{D.23}$$

whereby completing the square yields $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 = O_P(a_{nT})$. \square

The proof of Proposition 4 requires the following lemma which is proven at the end of this Appendix.

Lemma D.1 Under Assumptions 1–4, if $\check{\boldsymbol{\theta}}$ satisfies $\|\check{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2 = O_P(a_{nT})$ then $\hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}}) = \sigma_0^2 + O_P(\sqrt{Q}a_{nT})$, where $\check{\boldsymbol{\Lambda}} := \hat{\boldsymbol{\Lambda}}(\check{\boldsymbol{\theta}})$.

Proof of Proposition 4. Let $\boldsymbol{\gamma}^0$ be some $\boldsymbol{\gamma}$ which satisfies Assumptions 3.2 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 Lemma D.1 and Assumption 8.1. Hence,

$$\begin{aligned} \text{IC}^*(\boldsymbol{\gamma}^0) &= \hat{\sigma}^2(\boldsymbol{\gamma}^0) + \varrho_\rho Q^0 + \varrho_\beta K^0 \\ &= \sigma_0^2 + o_P(1), \end{aligned} \quad (\text{D.24})$$

as $Q^0 \varrho_\rho, K^0 \varrho_\beta \rightarrow 0$ by Assumption 8.2. Now, consider some $\boldsymbol{\gamma} \in \Gamma^+$ that produces an overfitted model. It is shown that, as $n, T \rightarrow \infty$,

$$\Pr\left(\inf_{\boldsymbol{\gamma} \in \Gamma^+} \text{IC}^*(\boldsymbol{\gamma}) > \text{IC}^*(\boldsymbol{\gamma}^0)\right) \rightarrow 1. \quad (\text{D.25})$$

Recalling that $\text{IC}^*(\boldsymbol{\gamma}) := \hat{\sigma}^2(\boldsymbol{\gamma}) + \varrho_\rho |\mathcal{S}_\rho(\boldsymbol{\gamma})| + \varrho_\beta |\mathcal{S}_\beta(\boldsymbol{\gamma})|$, (D.25) is equivalent to

$$\Pr\left(\inf_{\boldsymbol{\gamma} \in \Gamma^+} \{\hat{\sigma}^2(\boldsymbol{\gamma}) + \varrho_\rho |\mathcal{S}_\rho(\boldsymbol{\gamma})| + \varrho_\beta |\mathcal{S}_\beta(\boldsymbol{\gamma})|\} > \hat{\sigma}^2(\boldsymbol{\gamma}^0) + \varrho_\rho Q^0 + \varrho_\beta K^0\right) \rightarrow 1. \quad (\text{D.26})$$

Let $\boldsymbol{\gamma}^+ := \arg \inf_{\boldsymbol{\gamma} \in \Gamma^+} \text{IC}^*(\boldsymbol{\gamma})$. Then (D.26) gives

$$\begin{aligned} &\Pr\left((\sqrt{Q}a_{nT})^{-1}(\hat{\sigma}^2(\boldsymbol{\gamma}^+) - \hat{\sigma}^2(\boldsymbol{\gamma}^0)) + (\sqrt{Q}a_{nT})^{-1}\varrho_\rho(|\mathcal{S}_\rho(\boldsymbol{\gamma}^+)| - Q^0) \right. \\ &\quad \left. + (\sqrt{Q}a_{nT})^{-1}\varrho_\beta(|\mathcal{S}_\beta(\boldsymbol{\gamma}^+)| - K^0) > 0\right) \rightarrow 1. \end{aligned} \quad (\text{D.27})$$

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 $\|\hat{\boldsymbol{\theta}}^+ - \boldsymbol{\theta}^0\|_2 = O_P(a_{nT})$. Thus, using Lemma D.1, $\hat{\sigma}^2(\boldsymbol{\gamma}^+) - \hat{\sigma}^2(\boldsymbol{\gamma}^0) = O_P(\sqrt{Q}a_{nT})$, and so $(\sqrt{Q}a_{nT})^{-1}(\hat{\sigma}^2(\boldsymbol{\gamma}^+) - \hat{\sigma}^2(\boldsymbol{\gamma}^0)) = O_P(1)$. By Assumption 8.2, $(\sqrt{Q}a_{nT})^{-1}\varrho_\rho, (\sqrt{Q}a_{nT})^{-1}\varrho_\beta \rightarrow \infty$, and since either $|\mathcal{S}_\rho(\boldsymbol{\gamma}^+)| - Q^0 > 0$ or $|\mathcal{S}_\beta(\boldsymbol{\gamma}^+)| - K^0 > 0$, or both, then (D.27) holds as $n, T \rightarrow \infty$. Finally, in the case of an underfitted model, one of either $\mathcal{S}_\rho(\boldsymbol{\gamma}) \not\supseteq \mathcal{S}_{T,\rho}$ or $\mathcal{S}_\beta(\boldsymbol{\gamma}) \not\supseteq \mathcal{S}_{T,\beta}$ must be true. Then,

$$\inf_{\boldsymbol{\gamma} \in \Gamma^-} \text{IC}^*(\boldsymbol{\gamma}) > \inf_{\boldsymbol{\gamma} \in \Gamma^-} \hat{\sigma}^2(\boldsymbol{\gamma}) \xrightarrow{P} \sigma_-^2 > \sigma_0^2, \quad (\text{D.28})$$

using Assumption 8.3. Hence,

$$\Pr\left(\inf_{\boldsymbol{\gamma} \in \Gamma^-} \text{IC}^*(\boldsymbol{\gamma}) > \text{IC}^*(\boldsymbol{\gamma}^0)\right) \rightarrow 1. \quad (\text{D.29})$$

Combined, (D.25) and (D.29) establish the result. \square

Proof of Lemma D.1. Consider $\hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}})$. Substituting in true DGP yields

$$\begin{aligned} \hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}}) &= \frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}}) + \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{f}_t^0 + \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t)' \\ &\quad \times \mathbf{M}_{\check{\boldsymbol{\Lambda}}}(\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}}) + \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{f}_t^0 + \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t). \end{aligned} \quad (\text{D.30})$$

Expanding the above

$$\begin{aligned} \hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}}) &= \frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}}))' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}}) + \frac{2}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}}))' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{f}_t^0 \\ &\quad + \frac{2}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}}))' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t + \frac{2}{nT} \sum_{t=1}^T (\mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{f}_t^0)' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t \\ &\quad + \frac{1}{nT} \sum_{t=1}^T (\mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t)' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t + \frac{1}{nT} \sum_{t=1}^T (\mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{f}_t^0)' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{f}_t^0 \\ &=: l_1 + \dots + l_6. \end{aligned} \quad (\text{D.31})$$

Now,

$$|l_1| \leq \frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}})\|_2^2 = O_P(a_{nT}^2), \quad (\text{D.32})$$

$$|l_2| \leq \frac{2}{nT} \left(\sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}} \|\mathbf{S}(\check{\boldsymbol{\rho}})\|_2 \|\mathbf{S}^{-1}\|_2 \|\boldsymbol{\Lambda}^0\|_F \|\mathbf{F}^0\|_F = O_P(a_{nT}), \quad (\text{D.33})$$

$$|l_3| \leq \frac{2}{nT} \left(\sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \check{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}} \|\mathbf{S}(\check{\boldsymbol{\rho}})\|_2 \|\mathbf{S}^{-1}\|_2 \|\boldsymbol{\varepsilon}\|_F = O_P(a_{nT}), \quad (\text{D.34})$$

$$|l_4| \leq \frac{2R^0}{nT} \|\mathbf{S}(\check{\boldsymbol{\rho}})\|_2^2 \|\mathbf{S}^{-1}\|_2^2 \|\boldsymbol{\Lambda}^0\|_2 \|\mathbf{F}^0\|_2 \|\boldsymbol{\varepsilon}\|_2 = O_P(a_{nT}), \quad (\text{D.35})$$

using Lemmas A.2(iii), A.2(vi) and A.3(i), and

$$\begin{aligned} l_5 &= \frac{1}{nT} \text{tr}((\mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon})' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1}\boldsymbol{\varepsilon}) \\ &= \frac{1}{nT} \text{tr}(((\mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1} - \mathbf{I}_n)\boldsymbol{\varepsilon})' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} (\mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1} - \mathbf{I}_n)\boldsymbol{\varepsilon}) + \frac{2}{nT} \text{tr}(((\mathbf{S}(\check{\boldsymbol{\rho}})\mathbf{S}^{-1} - \mathbf{I}_n)\boldsymbol{\varepsilon})' \mathbf{M}_{\check{\boldsymbol{\Lambda}}} \boldsymbol{\varepsilon}) \\ &\quad + \frac{1}{nT} \text{tr}(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}) - \frac{1}{nT} \text{tr}(\boldsymbol{\varepsilon}'\mathbf{P}_{\check{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}) \\ &= O_P(Qa_{nT}^2) + O_P(\sqrt{Q}a_{nT}) + \sigma_0^2 + O_P(a_{nT}^2), \end{aligned} \quad (\text{D.36})$$

using Lemmas A.2(vi) and A.2(viii), as well as Lemma 9 in Yu et al. (2008). Similarly,

$$\begin{aligned}
l_6 &= \frac{1}{nT} \text{tr}((\mathbf{S}(\check{\rho})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \mathbf{S}(\check{\rho})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{F}^{0'}) \\
&= \frac{1}{nT} \text{tr}(((\mathbf{S}(\check{\rho})\mathbf{S}^{-1} - \mathbf{I}_n)\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} (\mathbf{S}(\check{\rho})\mathbf{S}^{-1} - \mathbf{I}_n)\boldsymbol{\Lambda}^0\mathbf{F}^{0'}) + \frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) \\
&\quad + \frac{2}{nT} \text{tr}(((\mathbf{S}(\check{\rho})\mathbf{S}^{-1} - \mathbf{I}_n)\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) \\
&= O_P(Qa_{nT}^2) + \frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) + O_P(\sqrt{Q}a_{nT}), \tag{D.37}
\end{aligned}$$

using Lemmas A.2(iii) and A.3(i). Combining (D.31) – (D.37) gives

$$\hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}}) = \sigma_0^2 + \frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) + O_P(\sqrt{Q}a_{nT}). \tag{D.38}$$

Now, let $\boldsymbol{\Lambda}^* := (\boldsymbol{\Lambda}^0, \mathbf{A})$ for some $\mathbf{A} \in \mathbb{R}^{n \times (R-R^0)}$ such that $\boldsymbol{\Lambda}^{*\prime} \boldsymbol{\Lambda}^*/n$ has rank R . Then, using similar steps to (D.32) – (D.37), it can be established that

$$\begin{aligned}
\hat{\sigma}^2(\check{\boldsymbol{\theta}}, \boldsymbol{\Lambda}^*) &= \sigma_0^2 + \frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\boldsymbol{\Lambda}^*} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) + O_P(\sqrt{Q}a_{nT}) \\
&= \sigma_0^2 + O_P(\sqrt{Q}a_{nT}), \tag{D.39}
\end{aligned}$$

because $\mathbf{M}_{\boldsymbol{\Lambda}^*} \boldsymbol{\Lambda}^0 = \mathbf{0}_{n \times R^0}$. By definition

$$\hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}}) \leq \hat{\sigma}^2(\check{\boldsymbol{\theta}}, \boldsymbol{\Lambda}^*), \tag{D.40}$$

or, equivalently,

$$\hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}}) - \hat{\sigma}^2(\check{\boldsymbol{\theta}}, \boldsymbol{\Lambda}^*) \leq 0. \tag{D.41}$$

Using (D.38) and (D.39) in (D.41) yields

$$\frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) + O_P(\sqrt{Q}a_{nT}) \leq 0, \tag{D.42}$$

whereby, because $(\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}$ is positive semidefinite, $\frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) \geq 0$ and thus

$$\frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0\mathbf{F}^{0'})' \mathbf{M}_{\check{\Lambda}} \boldsymbol{\Lambda}^0\mathbf{F}^{0'}) = O_P(\sqrt{Q}a_{nT}). \tag{D.43}$$

This gives

$$\hat{\sigma}^2(\check{\boldsymbol{\theta}}, \check{\boldsymbol{\Lambda}}) = \sigma_0^2 + O_P(\sqrt{Q}a_{nT}). \tag{D.44}$$

□

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 $n \times n$ positive definite matrix \mathbf{B} , $\det(\mathbf{B})^{\frac{1}{n}} \leq \frac{1}{n} \text{tr}(\mathbf{B})$ by the inequality of arithmetic and geometric means. This inequality is satisfied with equality if and only if all of the eigenvalues of \mathbf{B} are the same and, where \mathbf{B} is diagonalisable, this then implies $\mathbf{B} = b\mathbf{I}_n$ for some constant b . Given \mathbf{B} is positive definite, all of its eigenvalues are positive, whereby b must be strictly positive. \square

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

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

\square

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

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

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,

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

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

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

$$\begin{aligned}
&\leq \left(\sum_{j=1}^n |(\mathbf{W}_q)_{ij}| \right) \left(\sum_{j=1}^n |(\mathbf{W}_q)_{ij'}| \right) (\mathbb{E} [(x_{\kappa jt}^*)^2])^{\frac{1}{2}} (\mathbb{E} [(x_{\kappa j't}^*)^2])^{\frac{1}{2}} \\
&\leq c_1,
\end{aligned} \tag{E.1}$$

because the fourth moment of $x_{\kappa it}^*$ is uniformly bounded, and the weights matrix \mathbf{W}_q is UB uniformly over q . It then follows from Markov's inequality that $\|\mathbf{W}_q \boldsymbol{\mathcal{X}}_\kappa^*\|_F = O_P(\sqrt{nT})$. Next it is shown that the second moment of y_{it} is uniformly bounded over i and t .

$$\begin{aligned}
\mathbb{E}[y_{it}^2] &= \mathbb{E} \left[\left(\sum_{j=1}^n (\mathbf{S}^{-1})_{ij} \left(\sum_{k=1}^K \beta_k^0 x_{kjt} + \sum_{r=1}^{R^0} \lambda_{jr}^0 f_{tr}^0 + \varepsilon_{jt} \right) \right)^2 \right] \\
&=: \mathbb{E} \left[\left(\sum_{j=1}^n (\mathbf{S}^{-1})_{ij} (a_{jt} + b_{jt} + c_{jt}) \right)^2 \right] \\
&= \mathbb{E} \left[\sum_{j=1}^n \sum_{j'=1}^n (\mathbf{S}^{-1})_{ij} (\mathbf{S}^{-1})_{ij'} (a_{jt} + b_{jt} + c_{jt}) (a_{j't} + b_{j't} + c_{j't}) \right] \\
&=: \mathbb{E}[l_1 + \dots + l_6].
\end{aligned} \tag{E.2}$$

First,

$$\begin{aligned}
\mathbb{E}[|l_1|] &\leq \mathbb{E} \left[\sum_{j=1}^n \sum_{j'=1}^n |(\mathbf{S}^{-1})_{ij}| |(\mathbf{S}^{-1})_{ij'}| |a_{jt}| |a_{j't}| \right] \\
&= \mathbb{E} \left[\sum_{j=1}^n \sum_{j'=1}^n |(\mathbf{S}^{-1})_{ij}| |(\mathbf{S}^{-1})_{ij'}| \left| \sum_{k=1}^K \beta_k^0 x_{kjt} \right| \left| \sum_{k=1}^K \beta_k^0 x_{k j't} \right| \right] \\
&\leq \sum_{j=1}^n \sum_{j'=1}^n |(\mathbf{S}^{-1})_{ij}| |(\mathbf{S}^{-1})_{ij'}| \left(\mathbb{E} \left[\left| \sum_{k=1}^K \beta_k^0 x_{kjt} \right|^2 \right] \right)^{\frac{1}{2}} \left(\mathbb{E} \left[\left| \sum_{k'=1}^K \beta_{k'}^0 x_{k' j't} \right|^2 \right] \right)^{\frac{1}{2}} \\
&\leq c_2,
\end{aligned} \tag{E.3}$$

under Assumptions 2.3 and 2.4. Under Assumption 1.1, 2.3, 2.4 and 2.6, similar steps establish that $\mathbb{E}[l_2], \dots, \mathbb{E}[l_6]$ are also uniformly bounded. Thus using Markov's inequality it follows that $\|\boldsymbol{\mathcal{Y}}_{-1}\|_2 = O_P(\sqrt{nT})$, and by the same steps as in (E.1) $\mathbf{W}_q \boldsymbol{\mathcal{Y}}_{-1} = O_P(\sqrt{nT})$. Finally consider covariates of the form $\sum_{k=1}^K \beta_k^0 \mathbf{G}_q \boldsymbol{\mathcal{X}}_k$.

$$\mathbb{E} \left[\left(\sum_{k=1}^K \beta_k^0 (\mathbf{G}_q \boldsymbol{\mathcal{X}}_k)_{it} \right)^2 \right] = \mathbb{E} \left[\sum_{j=1}^n \sum_{j'=1}^n (\mathbf{G}_q)_{ij} (\mathbf{G}_q)_{ij'} \left(\sum_{k=1}^K \beta_k^0 x_{kjt} \right) \left(\sum_{k'=1}^K \beta_{k'}^0 x_{k' j't} \right) \right]$$

$$\begin{aligned}
&\leq \sum_{j=1}^n \sum_{j'=1}^n |(\mathbf{G}_q)_{ij}| |(\mathbf{G}_q)_{ij'}| \mathbb{E} \left[\left| \left(\sum_{k=1}^K \beta_k^0 x_{kjt} \right) \right| \left| \left(\sum_{k'=1}^K \beta_{k'}^0 x_{k'j't} \right) \right| \right] \\
&\leq \sum_{j=1}^n \sum_{j'=1}^n |(\mathbf{G}_q)_{ij}| |(\mathbf{G}_q)_{ij'}| \left(\mathbb{E} \left[\left| \sum_{k=1}^K \beta_k^0 x_{kjt} \right|^2 \right] \right)^{\frac{1}{2}} \left(\mathbb{E} \left[\left| \sum_{k'=1}^K \beta_{k'}^0 x_{k'j't} \right|^2 \right] \right)^{\frac{1}{2}} \\
&\leq c_3, \tag{E.4}
\end{aligned}$$

using Assumptions 2.3 and 2.4, which gives the result $\|\sum_{k=1}^K \beta_k^0 \mathbf{G}_q \boldsymbol{\mathcal{X}}_k\|_F = O_P(\sqrt{nT})$. \square

Proof of Lemma A.2(iii). Follows similar steps to the first part of the proof of Lemma A.2(ii) using Assumption 2.6. \square

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

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

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

Proof of Lemma A.2(vii).

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

using Assumption 2.4. \square

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

$$\begin{aligned}
\sum_{q=1}^Q |\rho_q^0 - \rho_q| \|\mathbf{W}_q \mathbf{S}^{-1}\|_2 &\leq \|\mathbf{S}^{-1}\|_2 \sum_{q=1}^Q |\rho_q^0 - \rho_q| \|\mathbf{W}_q\|_2 \tag{E.5} \\
&\leq \|\mathbf{S}^{-1}\|_2 \left(\sum_{q=1}^Q |\rho_q^0 - \rho_q|^2 \right)^{\frac{1}{2}} \left(\sum_{q=1}^Q \|\mathbf{W}_q\|_2^2 \right)^{\frac{1}{2}}
\end{aligned}$$

$$= \|\mathbf{S}^{-1}\|_2 \|\boldsymbol{\rho}^0 - \boldsymbol{\rho}\|_2 \sqrt{Q} \sqrt{\left(\max_{1 \leq q \leq Q} \|\mathbf{W}_q\|_2^2 \right)} = O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2),$$

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

Proof of Lemma A.3(i).

$$\begin{aligned} \frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta})\|_2^2 &= \frac{1}{nT} \sum_{t=1}^T \text{tr}(\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta})(\boldsymbol{\theta}^0 - \boldsymbol{\theta})' \mathbf{Z}_t') \\ &= (\boldsymbol{\theta}^0 - \boldsymbol{\theta})' \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}_t' \mathbf{Z}_t (\boldsymbol{\theta}^0 - \boldsymbol{\theta}) \\ &\leq \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2^2 \mu_1 \left(\frac{1}{nT} \sum_{t=1}^T \mathbf{Z}_t' \mathbf{Z}_t \right) \\ &= \|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2^2 O_P(1) = O_P(\|\boldsymbol{\theta}^0 - \boldsymbol{\theta}\|_2^2), \end{aligned}$$

where the last step uses Assumption 4.3. This implies the result. \square

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

$$\begin{aligned} \hat{\sigma}^2(\hat{\boldsymbol{\theta}}, \boldsymbol{\Lambda}) &= \frac{1}{nT} \sum_{t=1}^T (\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \mathbf{X}_t \boldsymbol{\beta}^0 + \mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \boldsymbol{\varepsilon}_t - \mathbf{X}_t \hat{\boldsymbol{\beta}})' \mathbf{M}_{\boldsymbol{\Lambda}} \\ &\quad \times (\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \mathbf{X}_t \boldsymbol{\beta}^0 + \mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \boldsymbol{\varepsilon}_t - \mathbf{X}_t \hat{\boldsymbol{\beta}}). \end{aligned} \quad (\text{E.6})$$

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

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

It is established in the detailed proof of Proposition 1 given in Appendix D that l_1, l_3, l_5 are $o_P(1)$. For l_2 ,

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

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

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

using similar steps to those for l_2 . For l_6 ,

$$\begin{aligned}
l_6 &\leq \frac{1}{nT} \|\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1}\|_2^2 \|\mathbf{M}_\Lambda\|_2 \|\boldsymbol{\Lambda}^0\|_2^2 \sum_{t=1}^T \|\mathbf{f}_t^0\|_2 \|\mathbf{f}_t^0\|_2 \\
&\leq \frac{1}{nT} \|\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1}\|_2^2 \|\mathbf{M}_\Lambda\|_2 \|\boldsymbol{\Lambda}^0\|_2^2 \left(\sum_{t=1}^T \|\mathbf{f}_t^0\|_2 \right) \\
&= \frac{1}{nT} \|\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1}\|_2^2 \|\mathbf{M}_\Lambda\|_2 \|\boldsymbol{\Lambda}^0\|_2^2 \|\mathbf{F}^0\|_F^2 \\
&= \frac{1}{nT} O_P(n) O_P(T) = O_P(1), \tag{E.10}
\end{aligned}$$

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

$$\hat{\sigma}^2(\hat{\boldsymbol{\theta}}, \boldsymbol{\Lambda}) = \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})' \mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1}) + \underbrace{O_P(1)}_{\geq 0} + O_P(a_{nT}). \tag{E.11}$$

Now, by Lemma A.1 $\sigma_0^2 \det((\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})' \mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})^{\frac{1}{n}} \leq \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})' \mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})$ and therefore a lower bound on (E.11) can be found where the above inequality is satisfied with

equality. This occurs when $(\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})'\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1} = c_1\mathbf{I}_n$, with $c_1 > 0$. Hence

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

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

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

since the matrices \mathbf{S}^{-1} and $\mathbf{S}(\hat{\boldsymbol{\rho}})$ are UB. As a result $0 < \frac{\sigma_0^2}{n}\text{tr}((\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1})'\mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{S}^{-1}) = O(1)$. It then follows that $\hat{\sigma}^{-2}(\hat{\boldsymbol{\theta}}, \boldsymbol{\Lambda}) = O_P(1)$ and is strictly positive w.p.a.1. This yields the result. \square

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 only concern the correct model, for notational convenience it is assumed in this section that all of the regressors are relevant, i.e., $K = K^0, Q = Q^0$ and $P = P^0$, which means $\mathbf{Z}_{(1)} = \mathbf{Z}$ and $\mathbf{Z}_{(1)} = \mathbf{Z}$.

Lemma F.1 *Under Assumptions 1–2,*

- (i) $\|\hat{\boldsymbol{\Lambda}}\|_2 = \sqrt{n}, \|\hat{\boldsymbol{\Lambda}}\|_F = \sqrt{Rn}$;
- (ii) $\|\mathbf{F}^{0'}\boldsymbol{\varepsilon}'\|_2 = O_P(\sqrt{nT}), \|\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\|_F = O_P(\sqrt{nT}), \|\mathbf{F}^{0'}\boldsymbol{\varepsilon}\boldsymbol{\Lambda}^0\|_2 = O_P(\sqrt{nT})$;
- (iii) $\|(\text{vec}(\mathbf{G}'_1\hat{\boldsymbol{\Lambda}}), \dots, \text{vec}(\mathbf{G}'_Q\hat{\boldsymbol{\Lambda}}))\|_2 = O(\sqrt{Qn})$;
- (iv) $\|(\text{vec}(\mathbf{G}_1\boldsymbol{\varepsilon}), \dots, \text{vec}(\mathbf{G}_Q\boldsymbol{\varepsilon}))\|_2 = O_P(\sqrt{QnT})$;
- (v) $\mathbb{E}[\sum_{p=1}^P \|\mathbf{z}_p - \bar{\mathbf{z}}_p\|_2^2] = O(P \max\{n, T\})$, where $\bar{\mathbf{z}}_p := \mathbb{E}_{\mathcal{D}}[\mathbf{z}_p]$;
- (vi) $\mathbb{E}[\sum_{p=1}^P \|\boldsymbol{\varepsilon}\bar{\mathbf{z}}_p'\|_F^2] = O(Pn^2T)$;
- (vii) $\mathbb{E}[\sum_{p=1}^P \|\mathbf{F}^{0'}\boldsymbol{\varepsilon}'\bar{\mathbf{z}}_p\|_F^2] = O(PnT^2), \mathbb{E}[\sum_{p=1}^P \|\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\bar{\mathbf{z}}_p'\|_F^2] = O(Pn^2T), \mathbb{E}[\sum_{p=1}^P \sum_{p'=1}^P \|\bar{\mathbf{z}}_p'\boldsymbol{\varepsilon}\bar{\mathbf{z}}_{p'}'\|_F^2] = O(P^2n^2T^2)$;
- (viii) $\|\frac{1}{T}\boldsymbol{\Lambda}^{0'}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\boldsymbol{\Lambda}^{0'}\boldsymbol{\Lambda}^0\|_F = O_P(\frac{n}{\sqrt{T}}), \|\frac{1}{n}\mathbf{F}^{0'}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\mathbf{F}^0 - \sigma_0^2\mathbf{F}^{0'}\mathbf{F}^0\|_F = O_P(\frac{T}{\sqrt{n}})$;
- (ix) $\mathbb{E}[\sum_{p=1}^P \|\frac{1}{T}\bar{\mathbf{z}}_p'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{\Lambda}^0 - \sigma_0^2\bar{\mathbf{z}}_p'\boldsymbol{\Lambda}^0\|_2^2] = O(Pn^2), \mathbb{E}[\sum_{p=1}^P \|\frac{1}{n}\bar{\mathbf{z}}_p'\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\mathbf{F}^0 - \sigma_0^2\bar{\mathbf{z}}_p'\mathbf{F}^0\|_2^2] = O(PT^2)$;
- (x) $\frac{1}{nT}\text{tr}(\boldsymbol{\varepsilon}'\mathbf{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\varepsilon}\mathbf{M}_{\mathbf{F}^0}) = \sigma_0^2 + O_P\left(\frac{1}{\min\{n, T\}}\right)$.

Lemma F.2 Under Assumptions 1–6,

- (i) $\frac{1}{\sqrt{n}}\|\hat{\Lambda}-\Lambda^0\mathbf{H}^*\|_2$ and $\frac{1}{\sqrt{T}}\|\hat{\mathbf{F}}'-\mathbf{H}^{*-1}\mathbf{F}^{0'}\|_2$ are $O_P(\sqrt{Q}\|\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}\|_2)+O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right)$, where $\mathbf{H}^* := \frac{1}{nT}\mathbf{F}^{0'}\mathbf{F}^0\Lambda^0\hat{\Lambda}\Pi^{-1}$ and $\hat{\mathbf{F}}' := \frac{1}{n}\sum_{t=1}^T\hat{\boldsymbol{\epsilon}}'_t\hat{\Lambda}$, with Π being a diagonal $R \times R$ matrix containing the largest R eigenvalues of $\frac{1}{nT}\sum_{t=1}^T\hat{\boldsymbol{\epsilon}}_t\hat{\boldsymbol{\epsilon}}'_t$ along its diagonal and $\hat{\boldsymbol{\epsilon}}_t := \mathbf{S}(\hat{\boldsymbol{\rho}})\mathbf{y}_t - \mathbf{X}_t\hat{\boldsymbol{\beta}}$;
- (ii) The matrix $\frac{1}{n}\Lambda^0\hat{\Lambda}$ converges in probability to an invertible matrix;
- (iii) $\frac{1}{nT}\|\mathbf{F}^{0'}\boldsymbol{\varepsilon}'\hat{\Lambda}\|_2 = O_P\left(\frac{1}{\sqrt{nT}}\right) + O_P\left(\frac{1}{T}\right) + O_P\left(\frac{\sqrt{Q}\|\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}\|_2}{\sqrt{T}}\right) + O_P\left(\frac{\sqrt{P}\|\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}\|_2}{\min\{\sqrt{nT},T\}}\right)$;
- (iv) $-\frac{1}{n^2T^2}\sum_{t=1}^T\sum_{\tau=1}^T\mathbf{Z}'_t\mathbf{M}_{\hat{\Lambda}}\boldsymbol{\varepsilon}'_{\tau}\hat{\Lambda}\boldsymbol{\Sigma}^*\mathbf{f}_t^0 = O_P\left(\frac{\sqrt{QP}\|\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}\|_2}{\min\{n,T\}}\right) + O_P\left(\frac{\sqrt{P}}{\min\{n^{1.5},T^{1.5}\}}\right)$;
- (v) $\|\mathbf{P}_{\hat{\Lambda}}-\mathbf{P}_{\Lambda^0}\|_2$ and $\|\mathbf{P}_{\mathbf{F}^0}-\mathbf{P}_{\hat{\mathbf{F}}}\|_2$ are $O_P(\sqrt{Q}\|\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}\|_2)+O_P\left(\frac{1}{\sqrt{\min\{n,T\}}}\right)$;
- (vi) $\|\hat{\mathbf{F}}'\|_2 = O_P(\sqrt{T})$;
- (vii) $\frac{1}{\sqrt{nT}}\mathbf{Z}^{*'}(\mathbf{M}_{\mathbf{F}^0}\otimes\mathbf{M}_{\hat{\Lambda}})\text{vec}(\boldsymbol{\varepsilon}) = \frac{1}{\sqrt{nT}}\mathbf{Z}^{*'}(\mathbf{M}_{\mathbf{F}^0}\otimes\mathbf{M}_{\Lambda^0})\text{vec}(\boldsymbol{\varepsilon})+o_P(1)$, where $\mathbf{Z}^* := (\mathbf{Z}_1^*, \dots, \mathbf{Z}_T^*)'$ with $\mathbf{Z}_t^* := (\mathbf{W}_1\mathbf{y}_t, \dots, \mathbf{W}_Q\mathbf{y}_t, \mathbf{X}_t)$;
- (viii) $\frac{1}{\sqrt{nT}}\mathbf{Z}'(\mathbf{M}_{\mathbf{F}^0}\otimes\mathbf{M}_{\Lambda^0})\text{vec}(\boldsymbol{\varepsilon}) = O_P(\sqrt{P})$;
- (ix) $\frac{1}{nT}\mathbb{E}\left[\|\mathbf{Z}'\mathbf{Z}-\mathbf{Z}'(\mathbf{M}_{\mathbf{F}^0}\otimes\mathbf{M}_{\Lambda^0})\mathbf{Z}\|_2\right] = o(1)$.

Lemma F.3 Under Assumptions 1–6,

- (i) $\mathbf{B}_1 = \mathbf{B}_1^* + O_P(Q^{1.5}\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0\|_2)$, where \mathbf{B}_1 and \mathbf{B}_1^* are $Q \times Q$ matrices with (q, q') -th element equal to $\frac{1}{n}\text{tr}(\mathbf{G}_q(\bar{\boldsymbol{\rho}})\mathbf{G}_{q'}(\bar{\boldsymbol{\rho}}))$ and $\frac{1}{n}\text{tr}(\mathbf{G}_q\mathbf{G}_{q'})$, respectively;
- (ii) $\mathbf{B}_2 = \mathbf{B}_2^* + o_P(1)$, where \mathbf{B}_2 and \mathbf{B}_2^* are $Q \times Q$ matrices with (q, q') -th element equal to $\frac{1}{nT}\text{tr}((\mathbf{G}_q\boldsymbol{\varepsilon})'\mathbf{G}_{q'}\boldsymbol{\varepsilon})$ and $\frac{\sigma_0^2}{n}\text{tr}(\mathbf{G}'_q\mathbf{G}_{q'})$, respectively;
- (iii)
$$\frac{1}{nT}\sum_{t=1}^T(\mathbf{Z}_t^*)'\mathbf{M}_{\hat{\Lambda}}\mathbf{Z}_t^* - \mathcal{H} = \frac{1}{nT}\mathbf{Z}'(\mathbf{M}_{\mathbf{F}^0}\otimes\mathbf{M}_{\Lambda^0})\mathbf{Z} + \begin{pmatrix} \mathbf{B}_2^* & \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} & \mathbf{0}_{K \times K} \end{pmatrix} + O_P\left(\sqrt{QP}\|\boldsymbol{\theta}^0-\hat{\boldsymbol{\theta}}\|_2\right) + O_P\left(\frac{P}{\sqrt{\min\{n,T\}}}\right),$$

where $\mathcal{H} := \frac{1}{nT}\mathbf{Z}^{*'}(\mathbf{P}_{\mathbf{F}^0}\otimes\mathbf{M}_{\hat{\Lambda}})\mathbf{Z}^*$, \mathbf{B}_2^* is defined in part (ii), and both \mathbf{Z}_t^* and \mathbf{Z}^* are defined in Lemma F.2(vii);

$$(iv) \quad \frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t^*)' \mathbf{M}_{\hat{\Lambda}} (\Lambda^0 \mathbf{f}_t^0 + \varepsilon_t) = \frac{1}{nT} \mathbf{Z}^{*'} (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\hat{\Lambda}}) \text{vec}(\varepsilon) + \mathfrak{H}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + \Delta_1 \\ + \left(\mathcal{O}_P(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathcal{O}_P\left(\frac{P}{\sqrt{\min\{n, T\}}}\right) \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0),$$

where Δ_1 is a term of order $\mathcal{O}_P\left(\frac{\sqrt{P}}{\min\{n^{1.5}, T^{1.5}\}}\right) + \mathcal{O}_P\left(\frac{P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n, T\}}\right) + \mathcal{O}_P\left(\frac{\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2}{\min\{\sqrt{nT}, T\}}\right) + \mathcal{O}_P\left(\frac{Q\sqrt{P} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^3}{\sqrt{T}}\right)$, \mathfrak{H} is defined in part (iii), and both \mathbf{Z}_t^* and \mathbf{Z}^* are defined in Lemma F.2(vii);

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

where \mathbf{b}_3^* is $Q \times 1$ with q -th element $\frac{\sigma_q^2}{n} \text{tr}(\mathbf{G}_q)$, $\boldsymbol{\kappa} := \frac{1}{nT} \mathbf{Z}' (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda^0}) \mathbf{Z} + \mathbf{B}_2^*$, \mathbf{B}_2^* is defined in part (ii), and Δ_2 has the same order as Δ_1 in part (iv);

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

Lemma F.4 Under Assumptions 1–6,

- (i) $\sum_{p=1}^P \|\sum_{k=1}^K (\hat{\beta}_k^0 - \beta_k) (\mathbf{G}_p(\hat{\boldsymbol{\rho}}) - \mathbf{G}_p) \boldsymbol{\chi}_k\|_F^2 = \mathcal{O}_P(Q^2 K n T \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^4)$;
- (ii) $\sum_{p=1}^P \|\sum_{k=1}^K (\hat{\beta}_k^0 - \beta_k) \mathbf{G}_p \boldsymbol{\chi}_k\|_F^2 = \mathcal{O}_P(Q K n T \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2)$;
- (iii) $\sum_{p=1}^P \|\sum_{k=1}^K \beta_k^0 (\mathbf{G}_p - \mathbf{G}_p(\hat{\boldsymbol{\rho}})) \boldsymbol{\chi}_k\|_F^2 = \mathcal{O}_P(Q^2 n T \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2)$;
- (iv) $\sum_{p=1}^P \|\sum_{k=1}^K (\hat{\beta}_k - \beta_k^0) \mathbf{G}_p(\hat{\boldsymbol{\rho}}) \boldsymbol{\chi}_k\|_F^2 = \mathcal{O}_P(Q K n T \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2^2)$;
- (v) $\sum_{p=1}^P \|\sum_{k=1}^K \beta_k^0 \mathbf{G}_p(\hat{\boldsymbol{\rho}}) \boldsymbol{\chi}_k\|_F^2 = \mathcal{O}_P(Q n T)$.

Proof of Lemma A.4. Consider the first order condition (A.13)

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

Evaluating (A.14) at $\hat{\boldsymbol{\theta}}, \hat{\Lambda}$,

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

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

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

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

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

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

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

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

$$\begin{aligned} &\left(\begin{pmatrix} \mathbf{B}_1^* & \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} & \mathbf{0}_{K \times K} \end{pmatrix} + \frac{1}{\hat{\sigma}^2} (\mathcal{B}_3 - \mathcal{H}) + \mathcal{O}_P(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathcal{O}_P\left(\frac{P}{\sqrt{\min\{n, T\}}}\right) \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\ &= \frac{1}{\hat{\sigma}^2} \frac{1}{\sqrt{nT}} \boldsymbol{\mathcal{Z}}' (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\hat{\boldsymbol{\Lambda}}}) \text{vec}(\boldsymbol{\varepsilon}) + \mathcal{B}_1 + \boldsymbol{\Delta}_1 - \frac{\partial \varrho(\boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\zeta})}{\partial \boldsymbol{\theta}}. \end{aligned} \quad (\text{F.6})$$

where \mathcal{H} is defined in Lemma F.3(iii) and \mathbf{B}_1^* is defined in Lemma F.3(i). Note, by Lemmas F.3(v) and F.1(x), $\frac{1}{\hat{\sigma}^2} = \frac{1}{\sigma_0^2} + \mathcal{O}_P\left(\frac{\sqrt{P}}{\min\{n, T\}}\right) + \mathcal{O}_P(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2)$, and also that $\|\frac{1}{\sqrt{nT}} \boldsymbol{\mathcal{Z}}' (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\boldsymbol{\Lambda}^0}) \boldsymbol{\mathcal{Z}}\|_2 \leq \frac{1}{\sqrt{nT}} \|\boldsymbol{\mathcal{Z}}\|_2^2 \|\mathbf{M}_{\mathbf{F}^0}\|_2 \|\mathbf{M}_{\boldsymbol{\Lambda}^0}\|_2 = \mathcal{O}_P(1)$ using Assumption 4.3.

Using these, and applying Lemma F.3(iii) to $\mathcal{B}_3 - \mathcal{H}$, (F.6) becomes

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

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

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

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

$$\begin{aligned}
& \frac{1}{\hat{\sigma}^2} \frac{1}{\sqrt{nT}} \mathcal{Z}^{*'}(M_{F^0} \otimes M_{\Lambda^0}) \text{vec}(\boldsymbol{\varepsilon}) \\
& = \frac{1}{\hat{\sigma}^2} \frac{1}{\sqrt{nT}} \mathcal{Z}'(M_{F^0} \otimes M_{\Lambda^0}) \text{vec}(\boldsymbol{\varepsilon}) + \frac{1}{\hat{\sigma}^2} \frac{1}{\sqrt{nT}} \begin{pmatrix} \text{tr} \left((\mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{F}^{0'})' M_{\Lambda^0} \boldsymbol{\varepsilon} M_{F^0} \right) \\ \vdots \\ \text{tr} \left((\mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{F}^{0'})' M_{\Lambda^0} \boldsymbol{\varepsilon} M_{F^0} \right) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \\
& \quad + \frac{1}{\hat{\sigma}^2} \frac{1}{\sqrt{nT}} \begin{pmatrix} \text{tr} \left((\mathbf{G}_1 \boldsymbol{\varepsilon})' M_{\Lambda^0} \boldsymbol{\varepsilon} M_{F^0} \right) \\ \vdots \\ \text{tr} \left((\mathbf{G}_Q \boldsymbol{\varepsilon})' M_{\Lambda^0} \boldsymbol{\varepsilon} M_{F^0} \right) \\ \mathbf{0}_{K \times 1} \end{pmatrix}. \tag{F.9}
\end{aligned}$$

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

$$\left(\frac{1}{\sigma_0^2} \frac{1}{nT} \mathcal{Z}'(M_{F^0} \otimes M_{\Lambda^0}) \mathcal{Z} + \begin{pmatrix} \mathbf{B}_1^* + \mathbf{B}_2^* & \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} & \mathbf{0}_{K \times K} \end{pmatrix} + \mathcal{O}_P(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) \right)$$

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

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) is $\mathcal{O}_P(1)$. In addition, using Lemmas F.3(v) and F.3(vi) again,

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

$$\begin{aligned}
& -\frac{2}{\sigma_0^2} \frac{1}{\sqrt{nT}} (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \left(\begin{pmatrix} \mathbf{b}_3^* \\ \mathbf{0}_{K \times 1} \end{pmatrix} + \mathbf{O}_P \left(\sqrt{\frac{Q}{nT}} \right) \right) \begin{pmatrix} T \text{tr}(\mathbf{G}_1) \\ \vdots \\ T \text{tr}(\mathbf{G}_Q) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \\
& -\frac{1}{\sigma_0^2} \frac{1}{\sqrt{nT}} (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \left(\boldsymbol{\kappa} + \mathbf{O}_P(\sqrt{Q}P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathbf{O}_P \left(\frac{P}{\sqrt{\min\{n, T\}}} \right) \right) (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) \begin{pmatrix} T \text{tr}(\mathbf{G}_1) \\ \vdots \\ T \text{tr}(\mathbf{G}_Q) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \\
& -\frac{1}{\sigma_0^2} \frac{1}{\sqrt{nT}} \Delta_2 \begin{pmatrix} T \text{tr}(\mathbf{G}_1) \\ \vdots \\ T \text{tr}(\mathbf{G}_Q) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \\
& = \frac{1}{\sigma_0^2} \frac{1}{\sqrt{nT}} \left(\begin{pmatrix} \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{F^0}) \\ \vdots \\ \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{F^0}) \\ \mathbf{0}_{K \times 1} \end{pmatrix} - \text{tr}(\boldsymbol{\varepsilon}' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{F^0}) \begin{pmatrix} \frac{1}{n} \text{tr}(\mathbf{G}_1) \\ \vdots \\ \frac{1}{n} \text{tr}(\mathbf{G}_Q) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \right) \\
& -\frac{2}{\sigma_0^2} \frac{1}{\sqrt{nT}} \begin{pmatrix} T \text{tr}(\mathbf{G}_1) \\ \vdots \\ T \text{tr}(\mathbf{G}_Q) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \begin{pmatrix} \frac{\sigma_0^2}{n} \text{tr}(\mathbf{G}_1) \\ \vdots \\ \frac{\sigma_0^2}{n} \text{tr}(\mathbf{G}_Q) \\ \mathbf{0}_{K \times 1} \end{pmatrix}' (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \mathbf{o}_P(1) \\
& + \left(\mathbf{O}_P(Q) + \mathbf{O}_P(\sqrt{Q} \sqrt{nT} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathbf{O}_P(QP \sqrt{nT} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2) \right) \\
& + \mathbf{O}_P \left(\frac{\sqrt{Q}P \sqrt{nT} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{\min\{n, T\}}} \right) (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}). \tag{F.11}
\end{aligned}$$

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

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

$$\begin{aligned}
& \left. + \mathcal{O}_P(\sqrt{QP}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathcal{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{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}} \left(\begin{array}{c} \text{tr}((\boldsymbol{G}_1 \boldsymbol{\varepsilon})' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots \\ \text{tr}((\boldsymbol{G}_Q \boldsymbol{\varepsilon})' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \mathbf{0}_{K \times 1} \end{array} \right) - \text{tr}(\boldsymbol{\varepsilon}' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) \begin{pmatrix} \frac{1}{n} \text{tr}(\boldsymbol{G}_1) \\ \vdots \\ \frac{1}{n} \text{tr}(\boldsymbol{G}_Q) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \right) \\
&- \sqrt{nT} \frac{\partial \varrho(\boldsymbol{\theta}, \boldsymbol{\gamma}, \zeta)}{\partial \boldsymbol{\theta}} + \mathcal{O}_P(1) \\
&= \frac{1}{\sigma_0^2} \frac{1}{\sqrt{nT}} \boldsymbol{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}((\boldsymbol{G}_1^* \boldsymbol{\varepsilon})' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \vdots \\ \text{tr}((\boldsymbol{G}_Q^* \boldsymbol{\varepsilon})' \boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \\
&- \sqrt{nT} \frac{\partial \varrho(\boldsymbol{\theta}, \boldsymbol{\gamma}, \zeta)}{\partial \boldsymbol{\theta}} + \mathcal{O}_P(1), \tag{F.12}
\end{aligned}$$

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

$$\left\| \sqrt{nT} \frac{\partial \varrho(\boldsymbol{\theta}, \boldsymbol{\gamma}, \zeta)}{\partial \boldsymbol{\theta}} \right\|_2 \leq \sqrt{PnT} \max\{\gamma_\rho, \gamma_\beta\} \mathcal{O}_P(1) = \mathcal{O}_P(1), \tag{F.13}$$

where the inequality in (F.13) follows the same steps as those used to obtain (D.20), and the final line follows under Assumption 6.6. Moreover, recalling the definition given in (16), notice that

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

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

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

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

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

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

$$\begin{aligned}
& \left(\mathbf{D} + \mathcal{O}_P(\sqrt{Q}P\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathcal{O}_P\left(\frac{P}{\sqrt{\min\{n, T\}}}\right) \right) \sqrt{nT}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
&= \mathcal{O}_P(\sqrt{P}) + \mathcal{O}_P(\sqrt{Q}) + \mathcal{O}_P(1).
\end{aligned} \tag{F.17}$$

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

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

□

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

$$\mathbf{D} - \hat{\mathbf{D}} = \frac{1}{\hat{\sigma}^2} \frac{1}{nT} (\mathbf{Z} - \hat{\mathbf{Z}})'(\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda^0}) \mathbf{Z} + \frac{1}{\hat{\sigma}^2} \frac{1}{nT} \hat{\mathbf{Z}}'(\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda^0})(\mathbf{Z} - \hat{\mathbf{Z}})$$

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

Note that, for $p = 1, \dots, Q$,

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

and $\mathbf{z}_p - \hat{\mathbf{z}}_p = \mathbf{0}_{n \times T}$ otherwise, and also that, for $p = 1, \dots, Q$,

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

and $\hat{\mathbf{z}}_p = \mathbf{z}_p$ otherwise. As such, using (F.20) and (F.21), terms $\mathbf{L}_1, \dots, \mathbf{L}_6$ in (F.19) can be expanded. Starting with \mathbf{L}_1 ,

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

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

$$\|\mathbf{L}_{1.1}\|_F \leq \left(\sum_{p=1}^P \|(\mathbf{z}_p - \hat{\mathbf{z}}_p)^{(1)}\|_F^2 \right)^{\frac{1}{2}} \left(\sum_{p'=1}^P \|\mathbf{z}_{p'}\|_F^2 \right)^{\frac{1}{2}}$$

$$= O_P(Q\sqrt{KPnT}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2), \quad (\text{F.23})$$

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

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

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

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

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

$$\|\mathbf{L}_{3.1}\|_F \leq \|\mathbf{P}_{\hat{\mathbf{F}}} - \mathbf{P}_{\mathbf{F}^0}\|_2 \left(\sum_{p=1}^P \|\hat{\mathbf{Z}}_p^{(1)}\|_F^2 \right)$$

$$= O_P(Q^{1.5}KnT\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^3) + O_P\left(\frac{QKnT\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2}{\sqrt{\min\{n, T\}}}\right), \quad (\text{F.27})$$

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

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

$$\begin{aligned} \|\mathbf{L}_{3.4}\|_F &\leq \|\mathbf{P}_{\hat{\mathbf{F}}} - \mathbf{P}_{\mathbf{F}^0}\|_2 \left(\sum_{p=1}^P \|\hat{\boldsymbol{\xi}}_p^{(2)}\|_F^2 \right) \\ &= O_P(Q^{1.5}nT\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{QnT}{\sqrt{\min\{n, T\}}}\right), \end{aligned} \quad (\text{F.29})$$

where the above uses Lemmas F.2(v), F.4(iv) and F.4(v). Thus $\|\mathbf{L}_3\|_F = O_P(Q^{1.5}KnT\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{QKnT}{\sqrt{\min\{n, T\}}}\right)$. It can be shown that $\|\mathbf{L}_4\|_F$ is of the same order. For term \mathbf{L}_5 , $\|\mathbf{L}_5\|_2 \leq |\sigma_0^{-2} - \hat{\sigma}^{-2}| \|\frac{1}{nT} \mathbf{Z}'(\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda^0}) \mathbf{Z}\|_2$. 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 4.3, $\|\frac{1}{nT} \mathbf{Z}'(\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda^0}) \mathbf{Z}\|_2 = O_P(1)$. Finally, let $\boldsymbol{\Omega} - \hat{\boldsymbol{\Omega}} =: \hat{\boldsymbol{\Omega}}^{(1)} + \dots + \hat{\boldsymbol{\Omega}}^{(6)}$ with elements $(\boldsymbol{\Omega}^{(1)})_{qq'} = \frac{1}{n} \text{tr}(\mathbf{G}_q(\mathbf{G}_{q'} - \mathbf{G}_{q'}(\hat{\boldsymbol{\rho}})))$, $(\boldsymbol{\Omega}^{(2)})_{qq'} = \frac{1}{n} \text{tr}((\mathbf{G}_q - \mathbf{G}_q(\hat{\boldsymbol{\rho}}))\mathbf{G}_{q'}(\hat{\boldsymbol{\rho}}))$, $(\boldsymbol{\Omega}^{(3)})_{qq'} = \frac{1}{n} \text{tr}(\mathbf{G}_q(\mathbf{G}'_{q'} - \mathbf{G}'_{q'}(\hat{\boldsymbol{\rho}})))$, $(\boldsymbol{\Omega}^{(4)})_{qq'} = \frac{1}{n} \text{tr}((\mathbf{G}_q - \mathbf{G}_q(\hat{\boldsymbol{\rho}}))\mathbf{G}'_{q'}(\hat{\boldsymbol{\rho}}))$, $(\boldsymbol{\Omega}^{(5)})_{qq'} = \frac{2}{n^2} \text{tr}(\mathbf{G}_q(\hat{\boldsymbol{\rho}}) - \mathbf{G}_q)\text{tr}(\mathbf{G}'_{q'}(\hat{\boldsymbol{\rho}}))$ and $(\boldsymbol{\Omega}^{(6)})_{qq'} = \frac{2}{n^2} \text{tr}(\mathbf{G}_q)\text{tr}(\mathbf{G}'_{q'}(\hat{\boldsymbol{\rho}}) - \mathbf{G}'_{q'})$. For the first of these terms,

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

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 $\|\hat{\boldsymbol{\Omega}}^{(2)}\|_F^2, \dots, \|\hat{\boldsymbol{\Omega}}^{(5)}\|_F^2$ have the same order, and therefore $\|\mathbf{L}_6\|_F = O_P(Q^{1.5}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2)$. Combining the above results for $\|\mathbf{L}_1\|_F, \dots, \|\mathbf{L}_6\|_F$ yields $\|\mathbf{D}^{-1} - \hat{\mathbf{D}}^{-1}\|_2 = O_P(Q^{1.5}P\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{QP}{\sqrt{\min\{n, T\}}}\right)$. \square

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

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

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

$$\begin{aligned} & \mathbb{E}_{\mathcal{D}} \left[(\text{tr}((\mathbf{G}_q^* \boldsymbol{\varepsilon})' \mathbf{P}_{\Lambda^0} \boldsymbol{\varepsilon}))^2 \right] \\ &= (\mathcal{M}_{\varepsilon}^4 - 3\sigma_0^4) T \sum_{i=1}^n (\mathbf{P}_{\Lambda^0} \mathbf{G}_q^*)_{ii}^2 + \sigma_0^4 (\text{tr}(\mathbf{I}_T \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*))^2 \\ & \quad + \sigma_0^4 \text{tr}((\mathbf{I}_T \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*) (\mathbf{I}_T \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*)') + \sigma_0^4 \text{tr}((\mathbf{I}_T \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*) (\mathbf{I}_T \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*)), \end{aligned} \quad (\text{F.32})$$

then

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

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

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

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

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

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

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

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

then

$$\begin{aligned} & \sum_{q=1}^Q \mathbb{E} \left[\mathbb{E}_{\mathcal{D}} \left[\left(\text{tr}((\mathbf{G}_q^* \boldsymbol{\varepsilon})' \mathbf{P}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0}) - \sigma_0^2 R^0 \text{tr}(\mathbf{P}_{\Lambda^0} \mathbf{G}_q^*) \right)^2 \right] \right] \\ &= \sum_{q=1}^Q \mathbb{E} \left[(\mathcal{M}_{\boldsymbol{\varepsilon}}^4 - 3\sigma_0^4) \sum_{i=1}^{nT} (\mathbf{P}_{\mathbf{F}^0} \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*)_{ii}^2 + \sigma_0^4 \text{tr}((\mathbf{P}_{\mathbf{F}^0} \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*) (\mathbf{P}_{\mathbf{F}^0} \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*)) \right. \\ & \quad \left. + \sigma_0^4 \text{tr}((\mathbf{P}_{\mathbf{F}^0} \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*) (\mathbf{P}_{\mathbf{F}^0} \otimes \mathbf{P}_{\Lambda^0} \mathbf{G}_q^*)') \right]. \end{aligned} \quad (\text{F.37})$$

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

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

□

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

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

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

$$\begin{aligned} & \mathbb{E}_{\mathcal{D}} \left[(\text{tr}((\mathbf{G}_q^* \boldsymbol{\varepsilon})' \mathbf{P}_{\Lambda^0} \boldsymbol{\varepsilon}))^2 \right] \\ &= (\mathcal{M}_{\boldsymbol{\varepsilon}}^4 - 3\sigma_0^4) \sum_{i=1}^{nT} (\mathbf{P}_{\mathbf{F}^0} \otimes (\mathbf{G}_q^*)')_{ii}^2 + \sigma_0^4 T^2 (\text{tr}(\mathbf{P}_{\Lambda^0} \mathbf{G}_q^*))^2 \\ & \quad + \sigma_0^4 \text{tr}((\mathbf{P}_{\mathbf{F}^0} \otimes (\mathbf{G}_q^*)') (\mathbf{P}_{\mathbf{F}^0} \otimes (\mathbf{G}_q^*)')) + \sigma_0^4 \text{tr}((\mathbf{P}_{\mathbf{F}^0} \otimes (\mathbf{G}_q^*)') (\mathbf{P}_{\mathbf{F}^0} \otimes (\mathbf{G}_q^*)')), \end{aligned} \quad (\text{F.40})$$

then

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

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

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

□

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

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

First,

$$\|\mathbf{l}_1\|_2^2 = \frac{1}{n^2} \sum_{p=1}^P \text{tr} \left(\left(\frac{1}{n} \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \right)^{-1} \boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} (\mathbf{Z}_p - \bar{\mathbf{Z}}_p)' \boldsymbol{\Lambda}^0 \right)^2$$

$$\leq \frac{(R^0)^2}{n^2} \left\| \frac{1}{n} \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \right\|_2^2 \sum_{p=1}^P \left\| \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon}(\mathbf{Z}_p - \bar{\mathbf{Z}}_p)' \mathbf{\Lambda}^0 \right\|_2^2. \quad (\text{F.44})$$

Now notice

$$\begin{aligned} & \mathbb{E} \left[\sum_{p=1}^P \left\| \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon}(\mathbf{Z}_p - \bar{\mathbf{Z}}_p)' \mathbf{\Lambda}^0 \right\|_2^2 \right] \\ & \leq \sum_{p=1}^P \mathbb{E} \left[\left\| \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon}(\mathbf{Z}_p - \bar{\mathbf{Z}}_p)' \mathbf{\Lambda}^0 \right\|_F^2 \right] \\ & = \sum_{p=1}^P \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \left(\sum_{i=1}^n \sum_{t=1}^T \sum_{j=1}^n \lambda_{ir}^0 \varepsilon_{it}(\mathbf{Z}_p - \bar{\mathbf{Z}}_p)_{jt} \lambda_{js}^0 \right)^2 \right] \\ & = \sum_{p=1}^P \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{i=1}^n \sum_{t=1}^T \sum_{j=1}^n \sum_{i'=1}^n \sum_{t'=1}^T \sum_{j'=1}^n \lambda_{ir}^0 \lambda_{js}^0 \lambda_{i'r}^0 \lambda_{j's}^0 \mathbb{E}_{\mathcal{D}} [\varepsilon_{it}(\mathbf{Z}_p - \bar{\mathbf{Z}}_p)_{jt} \varepsilon_{i't'}(\mathbf{Z}_p - \bar{\mathbf{Z}}_p)_{j't'}] \right]. \end{aligned} \quad (\text{F.45})$$

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

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

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 $\mathbf{Z}_p - \bar{\mathbf{Z}}_p = \mathbf{0}_{n \times T}$. Using this and (F.44) gives the result that $\|\mathbf{l}_1\|_2^2 = O_P(PT)$ whereby $\frac{1}{\sqrt{nT}} \|\mathbf{l}_1\|_2 = O_P(\sqrt{P/n}) = o_P(1)$. For term \mathbf{l}_2 ,

$$\begin{aligned} \|\mathbf{l}_2\|_2^2 &= \sum_{p=1}^P \text{tr} \left((\mathbf{Z}_p - \bar{\mathbf{Z}}_p)' \mathbf{P}_{\mathbf{\Lambda}^0} \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0} \right)^2 \\ &\leq (R^0)^2 \sum_{p=1}^P \|\mathbf{Z}_p - \bar{\mathbf{Z}}_p\|_2^2 \|\mathbf{P}_{\mathbf{\Lambda}^0} \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0}\|_2^2 \\ &\leq \frac{(R^0)^2}{n^2 T^2} \left(\sum_{p=1}^P \|\mathbf{Z}_p - \bar{\mathbf{Z}}_p\|_2^2 \right) \|\mathbf{\Lambda}^0\|_2^2 \|\mathbf{F}^0\|_2^2 \left\| \left(\frac{1}{n} \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \right)^{-1} \right\|_2^2 \left\| \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \right\|_2^2 \|\mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon} \mathbf{F}^0\|_2^2 \end{aligned}$$

$$= O_P(P \max\{n, T\}), \quad (\text{F.47})$$

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

$$\begin{aligned} \frac{1}{\sqrt{nT}} \mathbf{l}_3 &= \frac{1}{\sqrt{nT}} \begin{pmatrix} \text{tr}((\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0}) - \mathbb{E}_{\mathcal{D}}[\text{tr}((\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0})] \\ \vdots \\ \text{tr}((\mathbf{Z}_P - \bar{\mathbf{Z}}_P)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0}) - \mathbb{E}_{\mathcal{D}}[\text{tr}((\mathbf{Z}_P - \bar{\mathbf{Z}}_P)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0})] \end{pmatrix} \\ &+ \frac{1}{\sqrt{nT}} \begin{pmatrix} \mathbb{E}_{\mathcal{D}}[\text{tr}((\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0})] \\ \vdots \\ \mathbb{E}_{\mathcal{D}}[\text{tr}((\mathbf{Z}_P - \bar{\mathbf{Z}}_P)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0})] \end{pmatrix} \\ &:= \frac{1}{\sqrt{nT}} \mathbf{l}_{3.1} + \frac{1}{\sqrt{nT}} \begin{pmatrix} \mathbb{E}_{\mathcal{D}}[\text{tr}((\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0})] \\ \vdots \\ \mathbb{E}_{\mathcal{D}}[\text{tr}((\mathbf{Z}_P - \bar{\mathbf{Z}}_P)' \boldsymbol{\varepsilon} \mathbf{P}_{\mathbf{F}^0})] \end{pmatrix}. \end{aligned} \quad (\text{F.48})$$

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

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

which, combined with the previous parts, yields the result. \square

Proof of Lemma A.6. In this proof a central limit theorem is proven for \mathbf{Sc} . 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 $\mathbf{v} \in \mathbb{R}^L$ with elements v_l and $\|\mathbf{v}\|_2$ bounded for all

L. Also, recall the definition of \mathbf{S} given in Assumption 7.1. Then $\mathbf{v}'\mathbf{S}\mathbf{c}$ equals

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

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

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

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

$$\mathbf{B} := \sum_{l=1}^L \sum_{p=1}^P v_l(\mathbf{S})_{lp} \frac{1}{2} (\mathbf{g}_p + \mathbf{g}'_p) \tag{F.53}$$

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

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

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

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

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 \mathbf{v} . Following Yu et al. (2008), define the σ -algebra $\mathcal{F}_{it} := \sigma(\varepsilon_{11}, \dots, \varepsilon_{n1}, \varepsilon_{12}, \dots, \varepsilon_{n2}, \varepsilon_{1t}, \dots, \varepsilon_{it})$. A central limit theorem for martingale difference arrays applies to $\frac{\mathcal{J}}{\sigma_{\mathcal{J}}}$ if the following two conditions are met:

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

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

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

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

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

Hence

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

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

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

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

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

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

$$\mathbb{E} \left[\left((|u_{it-1}| + |d_{it}|)^p + \sum_{j=1}^i |b_{ij}| \right)^{\frac{q}{p}} \right] \leq 2^{\frac{q}{p}-1} \left(\mathbb{E} [(|u_{it-1}| + |d_{it}|)^q] + \left(\sum_{j=1}^i |b_{ij}| \right)^{\frac{q}{p}} \right)$$

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

Now,

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

Hence

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

Recall that $q = 2 + \delta$. By the same steps as in Lemma 10 in Yu et al. (2008), it can be shown that $\mathbb{E}[(| \sum_{h=1}^{\infty} \mathcal{A}_p \mathbf{A}^{h-1} \mathbf{S}^{-1} \boldsymbol{\varepsilon}_{h+1}^* |_{it})^4]$ 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,

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

Hence

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

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

$$\mathbb{E}[|j_{it}|^q] \leq \left(O(P^{1+\frac{\delta}{2}}) + O(P^2) + O(P^{\frac{q}{2p}}) \right) O(\sqrt{P})$$

$$= O(P^{\frac{3+\delta}{2}}), \quad (\text{F.64})$$

therefore

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

and

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

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

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

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

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

First,

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

$$+ 2 \left(\sum_{j=1}^{i-1} b_{ij} \varepsilon_{jt} \right) \varepsilon_{it} b_{ii} (\varepsilon_{it}^2 - \sigma_0^2) + 4 \varepsilon_{it}^2 \left(\sum_{j=1}^{i-1} b_{ij} \varepsilon_{jt} \right) \left(\sum_{j=1}^{i-1} b_{ij} \varepsilon_{jt} \right) | \mathcal{F}_{i-1t} \Big].$$

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 $i, j = 1, \dots, n$. Thus,

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

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

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

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

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

$$=: \frac{4\sigma_0^2}{\frac{1}{nT}\sigma_{\mathcal{J}}^2} (H_1 + 2H_2 + H_3). \quad (\text{F.71})$$

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,

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

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 $\mathbb{E}[2H_2^2]$,

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

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

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

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

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

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),

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

Rearrange l_1 to give

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

For the first term of (F.77),

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

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

$$\begin{aligned}
&\sigma_0^2 \sum_{t=1}^T \sum_{i=1}^n \mathbb{E}[(u_{it-1} + d_{it})^2] \\
&= \sigma_0^2 \sum_{t=1}^T \sum_{i=1}^n \mathbb{E} \left[\left(\sum_{l=1}^L \sum_{p=1}^P v_l(\mathbf{S})_{lp} (\mathbf{Z})_{pit} \right)^2 \right] \\
&= \sigma_0^2 \mathbf{v}' \mathbf{S} \mathbb{E} \left[\begin{pmatrix} \text{tr}(\mathbf{Z}'_1 \mathbf{Z}_1) & \cdots & \text{tr}(\mathbf{Z}'_1 \mathbf{Z}_P) \\ \vdots & \ddots & \vdots \\ \text{tr}(\mathbf{Z}'_P \mathbf{Z}_1) & \cdots & \text{tr}(\mathbf{Z}'_P \mathbf{Z}_P) \end{pmatrix} \right] \mathbf{S}' \mathbf{v} \\
&= \sigma_0^4 nT \mathbf{v}' \mathbf{S} \mathbb{E} \left[\begin{pmatrix} \frac{1}{\sigma_0^2} \frac{1}{nT} \begin{pmatrix} \text{tr}(\mathbf{Z}'_1 \mathbf{Z}_1) & \cdots & \text{tr}(\mathbf{Z}'_1 \mathbf{Z}_P) \\ \vdots & \ddots & \vdots \\ \text{tr}(\mathbf{Z}'_P \mathbf{Z}_1) & \cdots & \text{tr}(\mathbf{Z}'_P \mathbf{Z}_P) \end{pmatrix} - \frac{1}{\sigma_0^2} \frac{1}{nT} \mathbf{Z}' (\mathbf{M}_{F^0} \otimes \mathbf{M}_{\Lambda^0}) \mathbf{Z} \end{pmatrix} \right] \mathbf{S}' \mathbf{v} \\
&\quad + \sigma_0^4 nT \mathbf{v}' \mathbf{S} \mathbb{E} \left[\frac{1}{\sigma_0^2} \frac{1}{nT} \mathbf{Z}' (\mathbf{M}_{F^0} \otimes \mathbf{M}_{\Lambda^0}) \mathbf{Z} \right] \mathbf{S}' \mathbf{v} \\
&= \sigma_0^4 nT \mathbf{v}' \mathbf{S} \mathbb{E} \left[\frac{1}{\sigma_0^2} \frac{1}{nT} \mathbf{Z}' (\mathbf{M}_{F^0} \otimes \mathbf{M}_{\Lambda^0}) \mathbf{Z} \right] \mathbf{S}' \mathbf{v} + o(1), \tag{F.79}
\end{aligned}$$

using Lemma F.2(ix). Therefore

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

Next consider l_2 :

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

First, $l_{2,1}$ equals

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

$$= \mathcal{M}_\varepsilon^3 \mathbf{v}' \mathbf{S} \mathbb{E} [\mathbf{\Phi} + \mathbf{\Phi}'] \mathbf{S}' \mathbf{v}, \quad (\text{F.82})$$

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

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

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

$$l_2 = nT\sigma_0^4 \mathbf{v}' \mathbf{S} \mathbf{V} \mathbf{S}' \mathbf{v}. \quad (\text{F.84})$$

Bringing together these results,

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

and therefore,

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

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

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

This completes the proof. \square

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 $\mathcal{Z}_{(1)} = \mathcal{Z}$.

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

Proof of Lemma F.1(ii).

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

using Assumptions 1 and 2.6, from which the first result follows. The remaining results are established similarly. \square

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

$$\begin{aligned}
&\left\| \left(\text{vec}(\mathbf{G}'_1 \hat{\boldsymbol{\Lambda}}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\boldsymbol{\Lambda}}) \right) \right\|_2 \leq \left\| \left(\text{vec}(\mathbf{G}'_1 \hat{\boldsymbol{\Lambda}}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\boldsymbol{\Lambda}}) \right) \right\|_F = \sqrt{\sum_{q=1}^Q \|\text{vec}(\mathbf{G}'_q \hat{\boldsymbol{\Lambda}})\|_2^2} \\
&= \sqrt{\sum_{q=1}^Q \|\mathbf{G}'_q \hat{\boldsymbol{\Lambda}}\|_F^2} \leq \sqrt{\sum_{q=1}^Q \|\mathbf{G}_q\|_2^2 \|\hat{\boldsymbol{\Lambda}}\|_F^2} \leq \sqrt{R} \|\hat{\boldsymbol{\Lambda}}\|_2 \sqrt{Q} \sqrt{\max_{1 \leq q \leq Q} \|\mathbf{G}_q\|_2^2} = O(\sqrt{Qn}),
\end{aligned}$$

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

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). \square

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

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

Using these expressions, for $p = 1, \dots, Q$,

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

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

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

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, $\mathbb{E}[\|\mathbf{r}\|_2^2] \leq \mathbb{E}[\|\mathbf{r}\|_F^2] = O(n)$ follows from the proof of Lemma 4(2) in Shi and Lee (2017). Now,

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

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

Proof of Lemma F.1(vi).

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

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

Proof of Lemma F.1(vii).

$$\mathbb{E} \left[\sum_{p=1}^P \|\mathbf{F}^{0'} \boldsymbol{\varepsilon}' \bar{\mathbf{Z}}_p\|_F^2 \right] = \mathbb{E} \left[\sum_{p=1}^P \sum_{r=1}^{R^0} \sum_{t=1}^T \left(\sum_{\tau=1}^T \sum_{i=1}^n f_{\tau r}^0 \varepsilon_{i\tau} \bar{z}_{pit} \right)^2 \right]$$

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

This follows from Assumptions 1, 2.5 and 2.6 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).

$$\begin{aligned}
&\mathbb{E} \left[\left\| \frac{1}{T} \mathbf{\Lambda}^{0'} \varepsilon \varepsilon' \mathbf{\Lambda}^0 - \sigma_0^2 \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \right\|_F^2 \right] \\
&= \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \left(\frac{1}{T} \sum_{t=1}^T \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t - \sigma_0^2 \lambda_r^{0'} \lambda_s^0 \right)^2 \right] \\
&= \frac{1}{T^2} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{t=1}^T \sum_{\tau=1}^T \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t \varepsilon_{\tau}' \lambda_s^0 \lambda_r^{0'} \varepsilon_{\tau} \right] \\
&\quad - \frac{2}{T} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{t=1}^T \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t \lambda_r^{0'} \lambda_s^0 \right] + \sigma_0^4 \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \lambda_r^{0'} \lambda_s^0 \lambda_r^{0'} \lambda_s^0 \right] \\
&= \frac{1}{T^2} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{t=1}^T \sum_{\tau=1}^T \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t \varepsilon_{\tau}' \lambda_s^0 \lambda_r^{0'} \varepsilon_{\tau} \right] - \sigma_0^4 \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \lambda_r^{0'} \lambda_s^0 \lambda_r^{0'} \lambda_s^0 \right]. \tag{G.7}
\end{aligned}$$

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

$$\begin{aligned}
&\frac{1}{T^2} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{t=1}^T \sum_{\tau=1}^T \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t \varepsilon_{\tau}' \lambda_s^0 \lambda_r^{0'} \varepsilon_{\tau} \right] \\
&= \frac{1}{T^2} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{t=1}^T \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t \right] + \frac{1}{T^2} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{t=1}^T \sum_{\tau \neq t}^T \varepsilon_t' \lambda_s^0 \lambda_r^{0'} \varepsilon_t \varepsilon_{\tau}' \lambda_s^0 \lambda_r^{0'} \varepsilon_{\tau} \right] \\
&= \frac{1}{T} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} (\mathcal{M}_4^\varepsilon - 3\sigma_0^4) \sum_{i=1}^n (\lambda_s^0 \lambda_r^{0'})_{ii}^2 + \sigma_0^4 (\text{tr}(\lambda_s^0 \lambda_r^{0'})^2 + \text{tr}(\lambda_s^0 \lambda_r^{0'} \lambda_s^0 \lambda_r^{0'}) + \text{tr}(\lambda_s^0 \lambda_r^{0'} \lambda_r^0 \lambda_s^0)) \right]
\end{aligned}$$

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

Thus,

$$\begin{aligned}
& \mathbb{E} \left[\left\| \frac{1}{T} \boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \boldsymbol{\Lambda}^0 - \sigma_0^2 \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \right\|_F^2 \right] \\
& = \frac{1}{T} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{i=1}^n (\boldsymbol{\lambda}_s^0 \boldsymbol{\lambda}_r^{0'})_{ii}^2 + \sigma_0^4 (\boldsymbol{\lambda}_r^{0'} \boldsymbol{\lambda}_s^0 \boldsymbol{\lambda}_r^{0'} \boldsymbol{\lambda}_s^0 + \boldsymbol{\lambda}_s^{0'} \boldsymbol{\lambda}_s^0 \boldsymbol{\lambda}_r^{0'} \boldsymbol{\lambda}_r^0) \right] \\
& = \frac{(\mathcal{M}_4^\varepsilon - 3\sigma_0^4)}{T} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{i=1}^n \lambda_{i,r}^{02} \lambda_{i,s}^{02} \right] + \frac{\sigma_0^4}{T} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{i=1}^n \sum_{j=1}^n \lambda_{ir}^0 \lambda_{is}^0 \lambda_{jr}^0 \lambda_{js}^0 \right] \\
& + \frac{\sigma_0^4}{T} \mathbb{E} \left[\sum_{r=1}^{R^0} \sum_{s=1}^{R^0} \sum_{i=1}^n \sum_{j=1}^n \lambda_{is}^{02} \lambda_{jr}^{02} \right] = O\left(\frac{n^2}{T}\right), \tag{G.9}
\end{aligned}$$

using Assumptions 1 and 2.6. Hence, $\left\| \frac{1}{T} \boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \boldsymbol{\Lambda}^0 - \sigma_0^2 \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \right\|_F = O_P\left(\frac{n}{\sqrt{T}}\right)$. The second part of the lemma follows from analogous steps. \square

Proof of Lemma F.1(ix).

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

$$= \frac{1}{T^2} \mathbb{E} \left[\sum_{p=1}^P \sum_{t=1}^T \sum_{r=1}^{R^0} \sum_{\tau=1}^T \sum_{\tau'=1}^T \varepsilon'_\tau \lambda_r^0 \bar{\mathbf{Z}}'_{pt} \varepsilon_\tau \varepsilon'_{\tau'} \lambda_r^0 \bar{\mathbf{Z}}'_{pt} \varepsilon_{\tau'} \right] - \sigma_0^4 \mathbb{E} \left[\sum_{p=1}^P \sum_{t=1}^T \sum_{r=1}^{R^0} \bar{\mathbf{Z}}'_{pt} \lambda_r^0 \bar{\mathbf{Z}}'_{pt} \lambda_r^0 \right]. \quad (\text{G.10})$$

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

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

Thus,

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

which establishes the first part. The second part is obtained similarly. \square

Proof of Lemma F.1(x).

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

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

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

Similarly for l_3 and l_4 , which gives the result. \square

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

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

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 $\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} = \mathbf{I}_n + \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q$ by Lemma A.2(i), expand (G.15) as

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

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

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

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

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

Since $\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0$ and $\frac{1}{n} \Lambda^{0'} \hat{\Lambda}$ are both asymptotically invertible, by Assumption 6.3 and Lemma F.2(ii) respectively, let $\Sigma^* := \text{plim}_{n,T \rightarrow \infty} \left(\frac{1}{nT} \mathbf{F}^{0'} \mathbf{F}^0 \Lambda^{0'} \hat{\Lambda} \right)^{-1}$ whereby,

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

or

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

Now,

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

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 $\|\Sigma^*\|_2 = O_P(1)$. Using Lemmas A.3(i) and F.1(i),

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

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

$$\frac{1}{\sqrt{n}} \|\mathbf{P}_2\|_2 \leq \frac{1}{\sqrt{n}} \frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2 \|\Lambda^0 \mathbf{f}_t^0\|_2 \|\hat{\Lambda}\|_2$$

$$\begin{aligned}
&\leq \frac{1}{\sqrt{n}} \frac{1}{\sqrt{nT}} \left(\frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{t=1}^T \|\boldsymbol{\Lambda}^0 \mathbf{f}_t^0\|_2^2 \right)^{\frac{1}{2}} \|\hat{\boldsymbol{\Lambda}}\|_2 \\
&\leq \frac{1}{\sqrt{n}} \frac{1}{\sqrt{nT}} \left(\frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}} \|\boldsymbol{\Lambda}^0 \mathbf{F}^{0'}\|_F \|\hat{\boldsymbol{\Lambda}}\|_2 \\
&= O_P(\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2).
\end{aligned} \tag{G.20}$$

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

$$\begin{aligned}
\frac{1}{\sqrt{n}} \|\mathbf{P}_3\|_2 &\leq \frac{1}{\sqrt{n}} \frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2 \|\boldsymbol{\varepsilon}'_t \hat{\boldsymbol{\Lambda}}\|_2 \\
&\leq \frac{1}{\sqrt{n}} \frac{1}{\sqrt{nT}} \left(\frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{t=1}^T \|\boldsymbol{\varepsilon}'_t \hat{\boldsymbol{\Lambda}}\|_2^2 \right)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{n}} \frac{1}{\sqrt{nT}} \left(\frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}} \|\boldsymbol{\varepsilon}' \hat{\boldsymbol{\Lambda}}\|_F \\
&\leq \frac{1}{\sqrt{n}} \frac{1}{\sqrt{nT}} \left(\frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}} \|\boldsymbol{\varepsilon}\|_2 \|\hat{\boldsymbol{\Lambda}}\|_F \\
&= O_P \left(\frac{\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{\min\{n, T\}}} \right).
\end{aligned} \tag{G.21}$$

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

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

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

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

$$= O_P(\sqrt{Q}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2). \quad (\text{G.23})$$

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

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

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

$$\begin{aligned} \frac{1}{\sqrt{n}}\|\mathbf{P}_9\|_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\sum_{t=1}^T\|\boldsymbol{\Lambda}^0\mathbf{f}_t^0\|_2\left\|\sum_{q=1}^Q(\rho_q^0 - \hat{\rho}_q)\mathbf{G}_q\right\|_2\|\boldsymbol{\Lambda}^0\mathbf{f}_t^0\|_2\|\hat{\boldsymbol{\Lambda}}\|_2 \\ &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\left\|\sum_{q=1}^Q(\rho_q^0 - \hat{\rho}_q)\mathbf{G}_q\right\|_2\left(\sum_{t=1}^T\|\boldsymbol{\Lambda}^0\mathbf{f}_t^0\|_2^2\right)\|\hat{\boldsymbol{\Lambda}}\|_2 \\ &= \frac{1}{\sqrt{n}}\frac{1}{nT}\left\|\sum_{q=1}^Q(\rho_q^0 - \hat{\rho}_q)\mathbf{G}_q\right\|_2\|\boldsymbol{\Lambda}^0\mathbf{F}^{0'}\|_F^2\|\hat{\boldsymbol{\Lambda}}\|_2 \\ &= O_P(\sqrt{Q}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2). \end{aligned} \quad (\text{G.25})$$

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

$$\begin{aligned} \frac{1}{\sqrt{n}}\|\mathbf{P}_{10}\|_2 &\leq \frac{1}{\sqrt{n}}\frac{1}{nT}\|\boldsymbol{\Lambda}^0\mathbf{F}^{0'}\|_F\|\boldsymbol{\varepsilon}\|_2\left\|\sum_{q=1}^Q(\rho_q^0 - \hat{\rho}_q)\mathbf{G}_q\right\|_2\|\hat{\boldsymbol{\Lambda}}\|_2 \\ &= O_P\left(\frac{\sqrt{Q}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{\min\{n, T\}}}\right). \end{aligned} \quad (\text{G.26})$$

Using Lemma F.1(i),

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

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

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

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

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

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

$$\frac{1}{\sqrt{n}} \|\mathbf{P}_{20}\|_2 \leq \frac{1}{\sqrt{n}} \frac{1}{nT} \left\| \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \right\|_2^2 \|\boldsymbol{\Lambda}^0 \mathbf{F}^{0'}\|_F \|\boldsymbol{\varepsilon}\|_2 \|\hat{\boldsymbol{\Lambda}}\|_2 = O_P \left(\frac{Q \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2}{\sqrt{\min\{n, T\}}} \right). \quad (\text{G.30})$$

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

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

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

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

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

$$\begin{aligned} \frac{1}{\sqrt{T}} (\hat{\mathbf{F}}' - \mathbf{H}^{*-1} \mathbf{F}^{0'}) &= \frac{1}{n\sqrt{T}} \hat{\boldsymbol{\Lambda}}' \left(\sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p) \boldsymbol{z}_p \right) + \frac{1}{n\sqrt{T}} \hat{\boldsymbol{\Lambda}}' (\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} - \mathbf{I}_n) \boldsymbol{\Lambda}^0 \mathbf{F}^{0'} \\ &\quad + \frac{1}{n\sqrt{T}} \hat{\boldsymbol{\Lambda}}' (\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} - \mathbf{I}_n) \boldsymbol{\varepsilon} + \frac{1}{n\sqrt{T}} \hat{\boldsymbol{\Lambda}}' (\boldsymbol{\Lambda}^0 - \hat{\boldsymbol{\Lambda}} \mathbf{H}^{*-1}) \mathbf{F}^{0'} \\ &\quad + \frac{1}{n\sqrt{T}} \hat{\boldsymbol{\Lambda}}' \boldsymbol{\varepsilon} \end{aligned}$$

$$=: \mathbf{L}_1 + \dots + \mathbf{L}_5. \quad (\text{G.33})$$

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

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

using Lemma F.1(i) and Assumption 4.3. Therefore $\|\mathbf{L}_1\|_2 = O_P(\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2)$. Next,

$$\|\mathbf{L}_2\|_2 \leq \frac{1}{n\sqrt{T}} \|\hat{\mathbf{\Lambda}}\|_2 \|\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} - \mathbf{I}_n\|_2 \|\boldsymbol{\Lambda}^0 \mathbf{F}^{0'}\|_2 = O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2), \quad (\text{G.35})$$

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

$$\begin{aligned} \|\mathbf{L}_4\|_2 &\leq \frac{1}{n\sqrt{T}} \|\hat{\mathbf{\Lambda}}\|_2 \|\boldsymbol{\Lambda}^0 - \hat{\mathbf{\Lambda}} \mathbf{H}^{*-1}\|_2 \|\mathbf{F}^0\|_2 \\ &= O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right), \end{aligned} \quad (\text{G.37})$$

and

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

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

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

$$\mathcal{L}(\boldsymbol{\theta}^0, \hat{\mathbf{\Lambda}}) - \varrho(\boldsymbol{\theta}^0, \boldsymbol{\gamma}, \zeta) \leq \mathcal{L}(\hat{\boldsymbol{\theta}}, \hat{\mathbf{\Lambda}}) - \varrho(\hat{\boldsymbol{\theta}}, \boldsymbol{\gamma}, \zeta) \leq \mathcal{L}(\hat{\boldsymbol{\theta}}, \hat{\mathbf{\Lambda}}). \quad (\text{G.39})$$

By the same steps as those used to obtain (D.5) and (D.19) in the detailed proof of Proposition 1 given in Appendix D, it can be shown that

$$\frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2 + o_P(1)) + o_P(1) \leq \mathcal{L}(\boldsymbol{\theta}^0, \hat{\mathbf{\Lambda}}) - \varrho(\boldsymbol{\theta}^0, \boldsymbol{\gamma}, \zeta). \quad (\text{G.40})$$

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

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

Using similar steps to those for terms k_1, \dots, k_5 analysed in the detailed 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 $l_3 = \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1})' \mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1}) + o_P(1)$. Consider the remaining terms. Using Lemma A.2(iii),

$$\begin{aligned}
|l_4| &= \frac{2}{nT} \left| \text{tr}((\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \boldsymbol{\varepsilon})' \mathbf{M}_{\hat{\boldsymbol{\Lambda}}} \mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} \boldsymbol{\Lambda}^0 \mathbf{F}^{0'}) \right| \\
&\leq \frac{2R^0}{nT} \|\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1}\|_2^2 \|\boldsymbol{\varepsilon}\|_2 \|\boldsymbol{\Lambda}^0 \mathbf{F}^{0'}\|_F \\
&= \frac{2R^0}{nT} O_P(\sqrt{\max\{n, T\}}) O_P(\sqrt{nT}) = o_P(1). \tag{G.42}
\end{aligned}$$

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

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

For term l_6 ,

$$l_6 = \frac{1}{nT} \sum_{t=1}^T ((\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} - \mathbf{I}_n) \boldsymbol{\Lambda}^0 \mathbf{f}_t^0)' \mathbf{M}_{\hat{\boldsymbol{\Lambda}}} (\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} - \mathbf{I}_n) \boldsymbol{\Lambda}^0 \mathbf{f}_t^0$$

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

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

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

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

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

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

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

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

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

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

Proof of Lemma F.2(iii). Write

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

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

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

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

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

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

$$\begin{aligned} \|\mathbf{L}_{1.1.1}\|_2 &\leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{F}^0 \boldsymbol{\Lambda}^{0'} \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{H}^*\|_2 \left(\sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p)^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^P \|\mathbf{F}^{0'} \boldsymbol{\varepsilon}' \bar{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\ &\leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{F}^0\|_2 \|\boldsymbol{\Lambda}^0\|_2 \|\hat{\boldsymbol{\Lambda}}\|_2 \|\boldsymbol{\Sigma}^* \mathbf{H}^*\|_2 \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \left(\sum_{p=1}^P \|\mathbf{F}^{0'} \boldsymbol{\varepsilon}' \bar{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \end{aligned}$$

$$= O_P \left(\frac{\sqrt{P} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{nT}} \right), \quad (\text{G.51})$$

and

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

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

$$\begin{aligned} \mathbf{L}_{1.2} &= \frac{1}{nT} \frac{1}{nT} \sum_{\tau=1}^T \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0(\mathbf{Z}_\tau(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}))' \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{H}^* \\ &= \frac{1}{nT} \frac{1}{nT} \sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p) \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\Lambda}^0 \mathbf{F}^{0'} \boldsymbol{z}'_p \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{H}^*. \end{aligned} \quad (\text{G.53})$$

Then,

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

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

$$\mathbf{L}_{1.3} := \frac{1}{nT} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \mathbf{P}_8 \boldsymbol{\Sigma}^* \mathbf{H}^* = \frac{1}{nT} \frac{1}{nT} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\Lambda}^0 \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{H}^*, \quad (\text{G.55})$$

and

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

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

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

and therefore

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

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

$$\begin{aligned} \mathbf{L}_{1.5} &= \frac{1}{nT} \frac{1}{nT} \sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p) \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \mathbf{Z}_p' \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{H}^* \\ &= \frac{1}{nT} \frac{1}{nT} \sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p) \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \bar{\mathbf{Z}}_p' \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{H}^* \\ &\quad + \frac{1}{nT} \frac{1}{nT} \sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p) \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} (\mathbf{Z}_p - \bar{\mathbf{Z}}_p)' \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{H}^* \\ &=: \mathbf{L}_{1.5.1} + \mathbf{L}_{1.5.2}. \end{aligned} \quad (\text{G.58})$$

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

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

Similarly,

$$\|\mathbf{L}_{1.5.2}\|_2 \leq \frac{1}{nT} \frac{1}{nT} \|\hat{\boldsymbol{\Lambda}}\|_2 \|\boldsymbol{\Sigma}^*\|_2 \|\mathbf{H}^*\|_2 \|\mathbf{F}^{0'} \boldsymbol{\varepsilon}'\|_2 \|\boldsymbol{\varepsilon}\|_2 \left(\sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p)^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^P \|\mathbf{Z}_p - \bar{\mathbf{Z}}_p\|_2^2 \right)^{\frac{1}{2}}$$

$$\begin{aligned}
&= \frac{1}{nT} \frac{1}{nT} \|\hat{\Lambda}\|_2 \|\Sigma^*\|_2 \|\mathbf{H}^*\|_2 \|\mathbf{F}^{0'} \boldsymbol{\varepsilon}'\|_2 \|\boldsymbol{\varepsilon}\|_2 \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \left(\sum_{p=1}^P \|\boldsymbol{z}_p - \tilde{\boldsymbol{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\
&= O_P \left(\frac{\sqrt{P} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n\sqrt{T}, T^{1.5}\}} \right), \tag{G.60}
\end{aligned}$$

using Lemmas F.1(i), F.1(ii) and F.1(v). Hence $\|\mathbf{L}_{1.5}\|_2 = O_P \left(\frac{\sqrt{P} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{T} \right) + O_P \left(\frac{\sqrt{P} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n\sqrt{T}, T^{1.5}\}} \right)$.

Next for $\mathbf{L}_{1.6} := \frac{1}{nT} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \mathbf{P}_{12} \Sigma^* \mathbf{H}^* = \frac{1}{nT} \frac{1}{nT} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \mathbf{F}^0 \boldsymbol{\Lambda}^{0'} \hat{\Lambda} \Sigma^* \mathbf{H}^*$,

$$\|\mathbf{L}_{1.6}\|_2 \leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{F}^{0'} \boldsymbol{\varepsilon}\|_2^2 \|\boldsymbol{\Lambda}^0\|_2 \|\hat{\Lambda}\|_2 \|\Sigma^*\|_2 \|\mathbf{H}^*\|_2 = O_P \left(\frac{1}{T} \right), \tag{G.61}$$

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

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

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

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

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

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

□

Proof of Lemma F.2(iv). Decompose

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

as

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

Consider the first term

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

For the first of these,

$$\begin{aligned} \|\mathbf{l}_{1.1}\|_2 &\leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}\|_2 \|\Sigma^*\|_2 \|\varepsilon\|_2^2 \|\hat{\Lambda}\|_2 \sum_{t=1}^T \|\mathbf{Z}_t\|_2 \|\mathbf{f}_t^0\|_2 \\ &\leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}\|_2 \|\Sigma^*\|_2 \|\varepsilon\|_2^2 \|\hat{\Lambda}\|_2 \left(\sum_{t=1}^T \|\mathbf{Z}_t\|_2^2 \right)^{\frac{1}{2}} \|\mathbf{F}^0\|_F \\ &= O_P \left(\frac{\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n, T\}} \right) + O_P \left(\frac{\sqrt{P}}{\min\{n^{1.5}, T^{1.5}\}} \right), \end{aligned} \quad (\text{G.67})$$

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

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

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

$$\begin{aligned} \|\mathbf{l}_{1.2.1}\|_2 &\leq \frac{1}{n} \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t\|_2 \|\Lambda^0\|_2 \left\| \left(\frac{1}{n} \Lambda^{0'} \Lambda^0 \right)^{-1} \right\|_2 \|\Lambda^{0'} \varepsilon\|_2 \|\varepsilon\|_2 \|\hat{\Lambda} - \Lambda^0 \mathbf{H}^*\|_2 \\ &\quad \times \|\Sigma^*\|_2 \|\mathbf{f}_t^0\|_2 \\ &\leq \frac{1}{n} \frac{1}{nT} \frac{1}{nT} \|\Lambda^0\|_2 \left\| \left(\frac{1}{n} \Lambda^{0'} \Lambda^0 \right)^{-1} \right\|_2 \|\Lambda^{0'} \varepsilon\|_2 \|\varepsilon\|_2 \|\hat{\Lambda} - \Lambda^0 \mathbf{H}^*\|_2 \|\Sigma^*\|_2 \end{aligned}$$

$$\begin{aligned}
& \times \left(\sum_{t=1}^T \|\mathbf{Z}_t\|_2^2 \right)^{\frac{1}{2}} \|\mathbf{F}^0\|_F \\
& = O_P \left(\frac{\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n, \sqrt{nT}\}} \right) + O_P \left(\frac{\sqrt{P}}{\min\{n^{1.5}, \sqrt{nT}\}} \right). \tag{G.69}
\end{aligned}$$

Next,

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

where

$$\begin{aligned}
\|\mathbf{l}_{1.2.2.1}\|_2 & \leq \frac{1}{n^2} \frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t\|_2 \|\boldsymbol{\Lambda}^0\|_2 \left\| \left(\frac{1}{n} \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \right)^{-1} \right\|_2 \left\| \frac{1}{T} \boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \boldsymbol{\Lambda}^0 - \sigma_0^2 \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \right\|_2 \\
& \quad \times \|\mathbf{H}^*\|_2 \|\boldsymbol{\Sigma}^*\|_2 \|\mathbf{f}_t^0\|_2 \\
& \leq \frac{1}{n^2} \frac{1}{nT} \left(\sum_{t=1}^T \|\mathbf{Z}_t\|_2^2 \right)^{\frac{1}{2}} \|\boldsymbol{\Lambda}^0\|_2 \left\| \left(\frac{1}{n} \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \right)^{-1} \right\|_2 \left\| \frac{1}{T} \boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \boldsymbol{\Lambda}^0 - \sigma_0^2 \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \right\|_2 \\
& \quad \times \|\mathbf{H}^*\|_2 \|\boldsymbol{\Sigma}^*\|_2 \|\mathbf{F}^0\|_2 \\
& = O_P \left(\frac{\sqrt{P}}{n\sqrt{T}} \right) \tag{G.71}
\end{aligned}$$

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

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

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

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

where

$$\|\mathbf{l}_{2.1}\|_2 \leq \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \|\mathbf{Z}_t\|_2 \|\boldsymbol{\varepsilon}\|_2^2 \|\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0 \mathbf{H}^*\|_2 \|\boldsymbol{\Sigma}^*\|_2 \|\mathbf{f}_t^0\|_2$$

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

and hence

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

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

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

Now,

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

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

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

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

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

□

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

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

and

$$\begin{aligned} \left\| \frac{1}{n} \hat{\Lambda}' (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H}^*) \right\|_2 &\leq \frac{1}{n} \|\hat{\Lambda}\|_2 \|\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H}^*\|_2 \\ &= O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right). \end{aligned} \quad (\text{G.81})$$

Thus,

$$\begin{aligned} \frac{1}{n} \boldsymbol{\Lambda}^{0'} (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H}^*) &= \frac{1}{n} \boldsymbol{\Lambda}^{0'} \hat{\Lambda} - \frac{1}{n} \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \mathbf{H}^* \\ &= O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right), \end{aligned} \quad (\text{G.82})$$

and

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

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

$$\mathbf{I}_R - \frac{1}{n} \mathbf{H}^{*'} \boldsymbol{\Lambda}^{0'} \boldsymbol{\Lambda}^0 \mathbf{H}^* = O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right). \quad (\text{G.84})$$

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

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

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

$$\begin{aligned}
\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0} &= \frac{1}{n}(\hat{\Lambda} - \Lambda^0 \mathbf{H})(\hat{\Lambda} - \Lambda^0 \mathbf{H})' + \frac{1}{n}(\hat{\Lambda} - \Lambda^0 \mathbf{H}) \mathbf{H}' \Lambda^{0'} \\
&\quad + \frac{1}{n} \Lambda^0 \mathbf{H}(\hat{\Lambda} - \Lambda^0 \mathbf{H})' + \frac{1}{n} \Lambda^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \Lambda^{0'} \Lambda^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \Lambda^{0'} \\
&=: \mathbf{L}_1 + \mathbf{L}_2 + \mathbf{L}_3 + \mathbf{L}_4.
\end{aligned}$$

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

$$\begin{aligned}
\|\mathbf{L}_1\|_2 &\leq \frac{1}{n} \|\Lambda - \Lambda^0 \mathbf{H}^*\|_2^2 = O_P(Q \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2) + O_P\left(\frac{\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n, T\}}\right) + O_P\left(\frac{1}{\min\{n, T\}}\right) \\
\|\mathbf{L}_2\|_2 = \|\mathbf{L}_3\|_2 &\leq \frac{1}{\sqrt{n}} \|\Lambda - \Lambda^0 \mathbf{H}^*\|_2 \frac{1}{\sqrt{n}} \|\Lambda^0\|_2 \|\mathbf{H}^*\|_2 = O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right) \\
\|\mathbf{L}_4\|_2 &\leq \frac{1}{n} \|\Lambda^0\|_2^2 \|\mathbf{H}^*\|_2^2 \left\| \mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}^{*'} \Lambda^{0'} \Lambda^0 \mathbf{H}^* \right)^{-1} \right\|_2 = O_P(\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + O_P\left(\frac{1}{\sqrt{\min\{n, T\}}}\right).
\end{aligned}$$

The three results above gives

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

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

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

$$\frac{1}{\sqrt{T}} \hat{\mathbf{F}}' = \frac{1}{\sqrt{T}n} \hat{\Lambda}' \left(\sum_{p=1}^P (\theta_p^0 - \hat{\theta}_p) \mathbf{z}_p \right) + \frac{1}{\sqrt{T}n} \hat{\Lambda}' (\mathbf{S}(\hat{\boldsymbol{\rho}}) \mathbf{S}^{-1} - \mathbf{I}_n) \Lambda^0 \mathbf{F}^{0'}$$

$$\begin{aligned}
& + \frac{1}{\sqrt{T}n} \hat{\Lambda} (\mathbf{S}(\hat{\rho})\mathbf{S}^{-1} - \mathbf{I}_n) \boldsymbol{\varepsilon} + \frac{1}{\sqrt{T}n} \hat{\Lambda} \boldsymbol{\Lambda}^0 \mathbf{F}^{0'} + \frac{1}{\sqrt{T}n} \hat{\Lambda} \boldsymbol{\varepsilon} \\
& =: \mathbf{L}_1 + \mathbf{L}_2 + \mathbf{L}_3 + \mathbf{L}_4 + \mathbf{L}_5.
\end{aligned} \tag{G.87}$$

It was shown in the proof of Lemma F.2(i) that $\|\mathbf{L}_1\|_2, \|\mathbf{L}_2\|_2, \|\mathbf{L}_3\|_2, \|\mathbf{L}_5\|_2 = O_P(1)$. For \mathbf{L}_4 , $\|\mathbf{L}_4^*\|_2 \leq \frac{1}{\sqrt{T}n} \|\hat{\Lambda}\|_2 \|\boldsymbol{\Lambda}^0 \mathbf{F}^{0'}\|_2 = O_P(1)$, using Lemmas A.2(iii) and F.1(i), which gives the result. \square

Proof of Lemma F.2(vii).

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

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

$$\begin{aligned}
\mathbf{T}_1 & = \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr}(\mathbf{Z}'_1 (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H}) (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H})' \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \vdots \\ \text{tr}(\mathbf{Z}'_P (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H}) (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H})' \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \end{pmatrix} \\
& + \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr}(\mathbf{Z}'_1 (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H}) \mathbf{H}' \boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \vdots \\ \text{tr}(\mathbf{Z}'_P (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H}) \mathbf{H}' \boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \end{pmatrix} \\
& + \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr}(\mathbf{Z}'_1 \boldsymbol{\Lambda}^0 \mathbf{H} (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H})' \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \vdots \\ \text{tr}(\mathbf{Z}'_P \boldsymbol{\Lambda}^0 \mathbf{H} (\hat{\Lambda} - \boldsymbol{\Lambda}^0 \mathbf{H})' \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr} \left(\mathbf{Z}'_1 \Lambda^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \Lambda^{0'} \Lambda^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \Lambda^{0'} \varepsilon \mathbf{M}_{\mathbf{F}^0} \right) \\ \vdots \\ \text{tr} \left(\mathbf{Z}'_P \Lambda^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \Lambda^{0'} \Lambda^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \Lambda^{0'} \varepsilon \mathbf{M}_{\mathbf{F}^0} \right) \end{pmatrix} \\
& =: \mathbf{T}_{1.1.1} + \mathbf{T}_{1.2} + \mathbf{T}_{1.3} + \mathbf{T}_{1.4}. \tag{G.89}
\end{aligned}$$

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

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

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

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

Considering the four terms above,

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

$$\begin{aligned}
\|\mathbf{T}_{1.2.2}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{R}{n} \|\hat{\Lambda} - \Lambda^0 \mathbf{H}\|_2 \|\mathbf{H}\|_2 \|\Lambda^{0'} \boldsymbol{\varepsilon}\|_2 \left(\sum_{p=1}^P \|\mathbf{z}_p - \bar{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\
&= O_P \left(\frac{\sqrt{QP} \sqrt{\max\{n, T\}} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{n}} \right) + O_P \left(\frac{\sqrt{P} \sqrt{\max\{n, T\}}}{\min\{n, \sqrt{nT}\}} \right), \quad (\text{G.93})
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{T}_{1.2.3}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{R}{nT} \|\hat{\Lambda} - \Lambda^0 \mathbf{H}\|_2 \|\mathbf{H}\|_2 \|\Lambda^{0'} \boldsymbol{\varepsilon} \mathbf{F}^0\|_2 \left\| \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \right\|_2 \|\mathbf{F}^0\|_2 \left(\sum_{p=1}^P \|\bar{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\
&= O_P \left(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \right) + O_P \left(\sqrt{\frac{P}{\min\{n, T\}}} \right), \quad (\text{G.94})
\end{aligned}$$

and

$$\begin{aligned}
\|\mathbf{T}_{1.2.4}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{R}{nT} \|\hat{\Lambda} - \Lambda^0 \mathbf{H}\|_2 \|\mathbf{H}\|_2 \|\Lambda^{0'} \boldsymbol{\varepsilon} \mathbf{F}^0\|_2 \left\| \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \right\|_2 \|\mathbf{F}^0\|_2 \left(\sum_{p=1}^P \|\mathbf{z}_p - \bar{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\
&= O_P \left(\frac{\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{\min\{n, T\}}} \right) + O_P \left(\frac{\sqrt{P}}{\min\{n, T\}} \right), \quad (\text{G.95})
\end{aligned}$$

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

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

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

$$\|\mathbf{T}_{1.3.2}\|_2 \leq \frac{1}{\sqrt{nT}} \frac{R}{n} \|\mathbf{H}\|_2 \|\hat{\Lambda} - \Lambda^0 \mathbf{H}\|_2 \|\boldsymbol{\varepsilon}\|_2 \left(\sum_{p=1}^P \|\mathbf{z}_p - \bar{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \|\Lambda^0\|_2$$

$$= O_P \left(\frac{\sqrt{QP} \sqrt{\max\{n, T\}} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{\min\{n, T\}}} \right) + O_P \left(\frac{\sqrt{P} \sqrt{\max\{n, T\}}}{\min\{n, T\}} \right), \quad (\text{G.97})$$

and

$$\begin{aligned} \|\mathbf{T}_{1.3.4}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{R}{nT} \|\boldsymbol{\Lambda}^0\|_2 \|\mathbf{H}\|_2 \|\hat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0 \mathbf{H}\|_2^2 \|\boldsymbol{\varepsilon} \mathbf{F}^0\|_2 \left\| \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \right\|_2 \|\mathbf{F}^0\|_2 \left(\sum_{p=1}^P \|\mathbf{z}_p - \tilde{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\ &= O_P \left(\frac{\sqrt{QP} \sqrt{\max\{n, T\}} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{T}} \right) + O_P \left(\frac{\sqrt{P} \sqrt{\max\{n, T\}}}{\min\{\sqrt{nT}, T\}} \right), \end{aligned} \quad (\text{G.98})$$

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

$$\begin{aligned} \mathbf{T}_{1.3.1} &= \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr}(\tilde{\mathbf{z}}_1' \boldsymbol{\Lambda}^0 \mathbf{H} \mathbf{H}' (\hat{\boldsymbol{\Lambda}} \mathbf{H}^{-1} - \boldsymbol{\Lambda}^0)' \boldsymbol{\varepsilon}) \\ \vdots \\ \text{tr}(\tilde{\mathbf{z}}_P' \boldsymbol{\Lambda}^0 \mathbf{H} \mathbf{H}' (\hat{\boldsymbol{\Lambda}} \mathbf{H}^{-1} - \boldsymbol{\Lambda}^0)' \boldsymbol{\varepsilon}) \end{pmatrix} \\ &= \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr}(\tilde{\mathbf{z}}_1' \boldsymbol{\Lambda}^0 \mathbf{H} \mathbf{H}' (\mathbf{P}_2 + \mathbf{P}_6 + \mathbf{P}_8 + \mathbf{P}_9 + \mathbf{P}_{11} + \mathbf{P}_{12} + \mathbf{P}_{17})' \boldsymbol{\varepsilon}) \\ \vdots \\ \text{tr}(\tilde{\mathbf{z}}_P' \boldsymbol{\Lambda}^0 \mathbf{H} \mathbf{H}' (\mathbf{P}_2 + \mathbf{P}_6 + \mathbf{P}_8 + \mathbf{P}_9 + \mathbf{P}_{11} + \mathbf{P}_{12} + \mathbf{P}_{17})' \boldsymbol{\varepsilon}) \end{pmatrix} \\ &\quad + O_P \left(\frac{\sqrt{QP} \sqrt{\max\{n, T\}} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{\min\{n, T\}}} \right) + O_P \left(\frac{\sqrt{P} \sqrt{\max\{n, T\}}}{\min\{n, T\}} \right) \\ &=: \mathbf{T}_{1.3.1.1} + \dots + \mathbf{T}_{1.3.1.7} + O_P \left(Q \sqrt{P} \sqrt{\max\{n, T\}} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2 \right) \\ &\quad + O_P \left(\frac{\sqrt{QP} \sqrt{\max\{n, T\}} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{\min\{n, T\}}} \right) + O_P \left(\frac{\sqrt{P} \sqrt{\max\{n, T\}}}{\min\{n, T\}} \right). \end{aligned} \quad (\text{G.99})$$

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),

$$\begin{aligned} \|\mathbf{T}_{1.3.1.1}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{1}{n} \frac{R}{nT} \|\boldsymbol{\Lambda}^0\|_2^2 \|\mathbf{H}\|_2^2 \|\mathbf{F}^0\|_2 \|\hat{\boldsymbol{\Lambda}}\|_2 \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \left(\sum_{p=1}^P \sum_{p'=1}^P \|\tilde{\mathbf{z}}_{p'}' \boldsymbol{\varepsilon} \tilde{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\ &= O_P \left(P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \right). \end{aligned} \quad (\text{G.100})$$

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

$$\|\mathbf{T}_{1.3.1.2}\|_2 \leq \frac{1}{n} \frac{R}{nT} \|\boldsymbol{\Lambda}^0\|_2 \|\mathbf{H}\|_2^2 \|\hat{\boldsymbol{\Lambda}}\|_2 \|\mathbf{F}^0\|_F \left(\sum_{p=1}^P \|\boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \tilde{\mathbf{z}}_p\|_2^2 \right)^{\frac{1}{2}} \left(\frac{1}{nT} \sum_{t=1}^T \|\mathbf{z}_t(\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})\|_2^2 \right)^{\frac{1}{2}}$$

$$= O_P(\sqrt{P}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2). \quad (\text{G.101})$$

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

$$\begin{aligned} \|\mathbf{T}_{1.3.1.3}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{1}{n} \frac{R}{nT} \|\boldsymbol{\Lambda}^0\|_2 \|\mathbf{H}\|_2^2 \|\hat{\boldsymbol{\Lambda}}\|_2 \|\boldsymbol{\varepsilon} \mathbf{F}^0\|_2 \left(\sum_{p=1}^P \|\boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \bar{\boldsymbol{z}}_p\|_2^2 \right)^{\frac{1}{2}} \\ &= O_P \left(\sqrt{\frac{P}{T}} \right). \end{aligned} \quad (\text{G.102})$$

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

$$\begin{aligned} \|\mathbf{T}_{1.3.1.4}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{1}{n} \frac{R}{nT} \|\boldsymbol{\Lambda}^0\|_2^2 \|\mathbf{H}\|_2^2 \|\hat{\boldsymbol{\Lambda}}\|_2 \left\| \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \right\|_2 \|\mathbf{F}^0\|_2^2 \left(\sum_{p=1}^P \|\boldsymbol{\Lambda}^{0'} \boldsymbol{\varepsilon} \bar{\boldsymbol{z}}'_p\|_2^2 \right)^{\frac{1}{2}} \\ &= O_P(\sqrt{QP}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2). \end{aligned} \quad (\text{G.103})$$

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

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

where

$$\begin{aligned} \|\mathbf{T}_{1.3.1.5.1}\|_2^2 &\leq \frac{1}{\sqrt{nT}} \frac{1}{n} \frac{1}{nT} \|\boldsymbol{\Lambda}^0\|_2 \|\mathbf{H}\|_2 \|\hat{\boldsymbol{\Lambda}}\|_2 \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \left(\sum_{p'=1}^P \|\bar{\boldsymbol{z}}_{p'} \boldsymbol{\varepsilon}'\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^P \|\boldsymbol{\varepsilon} \bar{\boldsymbol{z}}'_p\|_2^2 \right)^{\frac{1}{2}} \\ &= O_P \left(\frac{P\sqrt{n}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{T}} \right), \end{aligned} \quad (\text{G.105})$$

and

$$\|\mathbf{T}_{1.3.1.5.2}\|_2 \leq \frac{1}{\sqrt{nT}} \frac{1}{n} \frac{R}{nT} \|\boldsymbol{\Lambda}^0\|_2 \|\mathbf{H}\|_2^2 \|\hat{\boldsymbol{\Lambda}}\|_2 \|\boldsymbol{\varepsilon}\|_2 \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \left(\sum_{p'=1}^P \|\boldsymbol{z}_{p'} - \bar{\boldsymbol{z}}_{p'}\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{p=1}^P \|\boldsymbol{\varepsilon} \bar{\boldsymbol{z}}'_p\|_2^2 \right)^{\frac{1}{2}}$$

$$= O_P \left(\frac{P\sqrt{n}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n, T\}} \right), \quad (\text{G.106})$$

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

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

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

$$\begin{aligned} \|\mathbf{T}_{1.3.1.7}\|_2 &\leq \frac{1}{\sqrt{nT}} \frac{1}{n} \frac{R}{nT} \|\boldsymbol{\Lambda}^0\|_2 \|\mathbf{H}\|_2^2 \|\hat{\boldsymbol{\Lambda}}\|_2 \|\mathbf{F}^{0'}\|_2^2 \|\boldsymbol{\rho}^0 - \hat{\boldsymbol{\rho}}\|_2 \left(\sum_{p=1}^P \sum_{q=1}^Q \|\boldsymbol{\Lambda}^{0'} \mathbf{G}'_q \boldsymbol{\varepsilon}' \bar{\mathbf{Z}}'_p\|_2^2 \right)^{\frac{1}{2}} \\ &= O_P \left(\frac{\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{T}} \right). \end{aligned} \quad (\text{G.108})$$

Combining all the above results gives

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

since $T/n \rightarrow c$ 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 $\mathbf{T}_{1.3.3}$ yield

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

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

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

$$\begin{aligned}
& + \frac{1}{\sqrt{nT}} \frac{\sigma_0^2}{nT} \begin{pmatrix} \text{tr} \left(\Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \mathbf{F}^{0'} \bar{\mathbf{z}}'_1 \right) \\ \vdots \\ \text{tr} \left(\Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \mathbf{F}^{0'} \bar{\mathbf{z}}'_p \right) \end{pmatrix} \\
& - \frac{1}{\sqrt{nT}} \frac{1}{nT} \frac{1}{T} \begin{pmatrix} \text{tr} \left(\bar{\mathbf{z}}'_1 \Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \left(\frac{1}{n} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \mathbf{F}^0 - \sigma_0^2 \mathbf{F}^{0'} \mathbf{F}^0 \right) \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \mathbf{F}^{0'} \right) \\ \vdots \\ \text{tr} \left(\bar{\mathbf{z}}'_p \Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \left(\frac{1}{n} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \mathbf{F}^0 - \sigma_0^2 \mathbf{F}^{0'} \mathbf{F}^0 \right) \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \mathbf{F}^{0'} \right) \end{pmatrix} \\
& - \frac{1}{\sqrt{nT}} \frac{\sigma_0^2}{nT} \begin{pmatrix} \text{tr} \left(\bar{\mathbf{z}}'_1 \Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \mathbf{F}^{0'} \mathbf{P}_{\mathbf{F}^0} \right) \\ \vdots \\ \text{tr} \left(\bar{\mathbf{z}}'_p \Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \mathbf{F}^{0'} \mathbf{P}_{\mathbf{F}^0} \right) \end{pmatrix} + o_P(1) \\
& = \frac{1}{\sqrt{nT}} \frac{1}{nT} \begin{pmatrix} \text{tr} \left(\Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \left(\frac{1}{n} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \bar{\mathbf{z}}'_1 - \sigma_0^2 \mathbf{F}^{0'} \bar{\mathbf{z}}'_1 \right) \right) \\ \vdots \\ \text{tr} \left(\Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \left(\frac{1}{n} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \bar{\mathbf{z}}'_p - \sigma_0^2 \mathbf{F}^{0'} \bar{\mathbf{z}}'_p \right) \right) \end{pmatrix} \\
& - \frac{1}{\sqrt{nT}} \frac{1}{nT} \frac{1}{T} \begin{pmatrix} \text{tr} \left(\bar{\mathbf{z}}'_1 \Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \left(\frac{1}{n} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \mathbf{F}^0 - \sigma_0^2 \mathbf{F}^{0'} \mathbf{F}^0 \right) \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \mathbf{F}^{0'} \right) \\ \vdots \\ \text{tr} \left(\bar{\mathbf{z}}'_p \Lambda^0 \mathbf{H} \mathbf{H}' \hat{\Lambda}' \Lambda^0 \left(\frac{1}{n} \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} \mathbf{F}^0 - \sigma_0^2 \mathbf{F}^{0'} \mathbf{F}^0 \right) \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \mathbf{F}^{0'} \right) \end{pmatrix} \\
& + o_P(1) \\
& =: \mathbf{a} + \mathbf{b} + o_P(1). \tag{G.111}
\end{aligned}$$

For terms \mathbf{a} and \mathbf{b} ,

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

and

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

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

$$\begin{aligned}
\mathbf{T}_{1.4} &= \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr} \left(\bar{\mathbf{z}}_1' \mathbf{\Lambda}^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon} \right) \\ \vdots \\ \text{tr} \left(\bar{\mathbf{z}}_P' \mathbf{\Lambda}^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon} \right) \end{pmatrix} \\
&+ \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr} \left((\mathbf{z}_1 - \bar{\mathbf{z}}_1)' \mathbf{\Lambda}^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon} \right) \\ \vdots \\ \text{tr} \left((\mathbf{z}_P - \bar{\mathbf{z}}_P)' \mathbf{\Lambda}^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon} \right) \end{pmatrix} \\
&- \frac{1}{\sqrt{nT}} \frac{1}{n} \begin{pmatrix} \text{tr} \left(\mathbf{z}'_1 \mathbf{\Lambda}^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon}_{\mathbf{F}^0} \right) \\ \vdots \\ \text{tr} \left(\mathbf{z}'_P \mathbf{\Lambda}^0 \mathbf{H} \left(\mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \mathbf{H} \right)^{-1} \right) \mathbf{H}' \mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon}_{\mathbf{F}^0} \right) \end{pmatrix} \\
&=: \mathbf{T}_{1.4.1} + \mathbf{T}_{1.4.2} + \mathbf{T}_{1.4.3}. \tag{G.114}
\end{aligned}$$

First,

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

Second,

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

Third,

$$\|\mathbf{T}_{1.4.3}\|_2 \leq \frac{R}{\sqrt{nT}} \frac{1}{nT} \left(\sum_{p=1}^P \|\mathbf{z}_p\|_2^2 \right)^{\frac{1}{2}} \|\mathbf{\Lambda}^0\|_2 \|\mathbf{H}\|_2^2 \left\| \mathbf{I}_R - \left(\frac{1}{n} \mathbf{H}' \mathbf{\Lambda}^{0'} \mathbf{\Lambda}^0 \mathbf{H} \right)^{-1} \right\|_2 \|\mathbf{\Lambda}^{0'} \boldsymbol{\varepsilon}_{\mathbf{F}^0}\|_2$$

$$\begin{aligned}
& \times \left\| \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right) \right\|_2 \|\mathbf{F}^0\|_2 \\
& = O_P \left(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 \right) + O_P \left(\sqrt{\frac{P}{\min\{n, T\}}} \right). \tag{G.117}
\end{aligned}$$

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

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

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

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

□

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

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

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

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

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

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

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

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

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

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

as

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

using Lemmas A.2(iii) and F.1(ii). Combining all the above and using Assumption 6.4,

$$\frac{1}{\sqrt{nT}} \mathbf{Z}' (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda^0}) \text{vec}(\boldsymbol{\varepsilon}) = O_P(\sqrt{P}).\tag{G.125}$$

□

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

$$\text{tr}(\mathbf{M}_{\Lambda^0} \bar{\mathbf{z}}_p \mathbf{M}_{\mathbf{F}^0} + (\mathbf{z}_p - \bar{\mathbf{z}}_p)' (\mathbf{M}_{\Lambda^0} \bar{\mathbf{z}}_{p'} \mathbf{M}_{\mathbf{F}^0} + (\mathbf{z}_{p'} - \bar{\mathbf{z}}_{p'}))) - \text{tr}(\mathbf{M}_{\Lambda^0} \mathbf{z}_{p'} \mathbf{M}_{\mathbf{F}^0} \mathbf{z}_p')$$

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

Thus,

$$\begin{aligned}
&\mathbf{Z}'\mathbf{Z} - \mathbf{Z}'(\mathbf{M}_{F^0} \otimes \mathbf{M}_{\Lambda^0})\mathbf{Z} \\
&= (\text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)), \dots, \text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_P - \bar{\mathbf{Z}}_P)))' (\text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)), \dots, \text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_P - \bar{\mathbf{Z}}_P))) \\
&\quad + (\text{vec}((\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)\mathbf{P}_{F^0}), \dots, \text{vec}((\mathbf{Z}_P - \bar{\mathbf{Z}}_P)\mathbf{P}_{F^0}))' (\text{vec}((\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)\mathbf{P}_{F^0}), \dots, \text{vec}((\mathbf{Z}_P - \bar{\mathbf{Z}}_P)\mathbf{P}_{F^0})) \\
&\quad - (\text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)\mathbf{P}_{F^0}), \dots, \text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_P - \bar{\mathbf{Z}}_P)\mathbf{P}_{F^0}))' \\
&\quad \times (\text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_1 - \bar{\mathbf{Z}}_1)\mathbf{P}_{F^0}), \dots, \text{vec}(\mathbf{P}_{\Lambda^0}(\mathbf{Z}_P - \bar{\mathbf{Z}}_P)\mathbf{P}_{F^0})) \\
&=: \mathbf{L}_1 + \mathbf{L}_2 + \mathbf{L}_3,
\end{aligned} \tag{G.127}$$

and hence

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

Consider the first term in (G.128)

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

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

Proof of Lemma F.3(i).

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

$$=: \mathbf{B}^{**} + \mathbf{B}_1^*.$$

First note that, by adding and subtracting,

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

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

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

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

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

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

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

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

Therefore (G.132) becomes

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

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

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

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

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

$$\begin{aligned} \mathbf{B}_3 - \mathcal{H} &= \frac{1}{nT} \begin{pmatrix} \text{tr}(\boldsymbol{\mathcal{Z}}'_1 (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_1 \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}(\boldsymbol{\mathcal{Z}}'_1 (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_P \mathbf{M}_{\mathbf{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}(\boldsymbol{\mathcal{Z}}'_P (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_1 \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}(\boldsymbol{\mathcal{Z}}'_P (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_P \mathbf{M}_{\mathbf{F}^0}) \end{pmatrix} \\ &+ \frac{1}{nT} \begin{pmatrix} \left(\begin{array}{ccc} \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_1 \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_1 \mathbf{M}_{\mathbf{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_P \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_P \mathbf{M}_{\mathbf{F}^0}) \end{array} \right) \mathbf{0}_{P \times K} \\ \left(\begin{array}{ccc} \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_1 \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_1 \mathbf{M}_{\mathbf{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_P \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \boldsymbol{\mathcal{Z}}_P \mathbf{M}_{\mathbf{F}^0}) \end{array} \right) \mathbf{0}_{P \times K} \end{pmatrix}' \\ &+ \frac{1}{nT} \begin{pmatrix} \left(\begin{array}{ccc} \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \mathbf{G}_1 \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \mathbf{G}_Q \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \mathbf{G}_1 \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) & \cdots & \text{tr}(\mathbf{G}_Q \boldsymbol{\varepsilon} (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \mathbf{G}_Q \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \end{array} \right) \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} \quad \mathbf{0}_{K \times K} \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{nT} \begin{pmatrix} \text{tr}((\mathbf{W}_1 \mathbf{Y})' M_{\Lambda^0} \mathbf{W}_1 \mathbf{Y} M_{F^0}) & \cdots & \text{tr}((\mathbf{W}_1 \mathbf{Y})' M_{\Lambda^0} \mathcal{X}_K M_{F^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathcal{X}_K)' M_{\Lambda^0} \mathbf{W}_1 \mathbf{Y} M_{F^0}) & \cdots & \text{tr}(\mathcal{X}'_K M_{\Lambda^0} \mathcal{X}_K M_{F^0}). \end{pmatrix} \\
& =: \mathbf{L}_1 + \mathbf{L}_2 + \mathbf{L}_3 + \mathbf{L}_4 + \mathbf{L}_5. \tag{G.134}
\end{aligned}$$

For the first term,

$$\begin{aligned}
\|\mathbf{L}_1\|_F^2 &= \frac{1}{n^2 T^2} \sum_{p=1}^P \sum_{p'=1}^P \text{tr}(\mathcal{Z}'_p (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}) \mathcal{Z}_{p'} M_{F^0})^2 \\
&\leq \frac{1}{n^2 T^2} \sum_{p=1}^P \sum_{p'=1}^P \|\mathcal{Z}'_p (\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0})\|_F^2 \|\mathcal{Z}_{p'} M_{F^0}\|_F^2 \\
&\leq \frac{1}{n^2 T^2} \left(\sum_{p=1}^P \|\mathcal{Z}_p\|_F^2 \right) \left(\sum_{p'=1}^P \|\mathcal{Z}_{p'}\|_F^2 \right) \|\mathbf{P}_{\hat{\Lambda}} - \mathbf{P}_{\Lambda^0}\|_2^2 \|M_{F^0}\|_2^2. \tag{G.135}
\end{aligned}$$

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

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

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

$$\begin{aligned}
\mathbf{L}_{5.3} &= \frac{1}{nT} \begin{pmatrix} \left(\begin{array}{ccc} \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{G}_1\boldsymbol{\varepsilon}) & \cdots & \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{G}_Q\boldsymbol{\varepsilon}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{G}_1\boldsymbol{\varepsilon}) & \cdots & \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{G}_Q\boldsymbol{\varepsilon}) \end{array} \right) & \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} & & \mathbf{0}_{K \times K} \end{pmatrix} \\
&- \frac{1}{nT} \begin{pmatrix} \left(\begin{array}{ccc} \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_1\boldsymbol{\varepsilon}) & \cdots & \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_Q\boldsymbol{\varepsilon}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_1\boldsymbol{\varepsilon}) & \cdots & \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_Q\boldsymbol{\varepsilon}) \end{array} \right) & \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} & & \mathbf{0}_{K \times K} \end{pmatrix} \\
&- \frac{1}{nT} \begin{pmatrix} \left(\begin{array}{ccc} \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{G}_1\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) & \cdots & \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{G}_Q\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{G}_1\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) & \cdots & \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{G}_Q\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) \end{array} \right) & \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} & & \mathbf{0}_{K \times K} \end{pmatrix} \\
&+ \frac{1}{nT} \begin{pmatrix} \left(\begin{array}{ccc} \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_1\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) & \cdots & \text{tr}((\mathbf{G}_1\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_Q\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) \\ \vdots & \ddots & \vdots \\ \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_1\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) & \cdots & \text{tr}((\mathbf{G}_Q\boldsymbol{\varepsilon})'\mathbf{P}_{\Lambda^0}\mathbf{G}_Q\boldsymbol{\varepsilon}\mathbf{P}_{F^0}) \end{array} \right) & \mathbf{0}_{Q \times K} \\ \mathbf{0}_{K \times Q} & & \mathbf{0}_{K \times K} \end{pmatrix} \\
&=: \frac{1}{nT} (\mathbf{L}_{5.3.1} + \mathbf{L}_{5.3.2} + \mathbf{L}_{5.3.3} + \mathbf{L}_{5.3.4}). \tag{G.137}
\end{aligned}$$

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

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

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

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

□

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

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

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

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

where $\boldsymbol{\Sigma}^* := \left(\frac{1}{nT} \mathbf{F}^{0'} \mathbf{F}^0 \Lambda^0 \hat{\Lambda}\right)^{-1}$ 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,

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

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

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

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

For term $\mathcal{B}_{4.1.2}$,

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

Letting $\varpi_{\tau t}^0 := \mathbf{f}_{\tau}^{0'} \left(\frac{1}{T} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \mathbf{f}_t^0$,

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

Next

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

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

$$\left\| \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \sum_{\tau=1}^T \boldsymbol{\varepsilon}'_{\tau} \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{f}_t^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\boldsymbol{\Lambda}}} \mathbf{Z}_{\tau} \right\|_2$$

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

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

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

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

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

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

$$\mathcal{B}_{4.1.5} = \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \left(-\frac{1}{nT} \sum_{\tau=1}^T \mathbf{Z}_\tau (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) \left(\sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \varepsilon_\tau \right)' \hat{\Lambda} \Sigma^* \right) \mathbf{f}_t^0$$

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

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

$$\begin{aligned} & \left\| \frac{1}{nT} \frac{1}{nT} \sum_{\tau=1}^T \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{Z}_\tau \left(\sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \boldsymbol{\varepsilon}_\tau \right)' \hat{\Lambda} \boldsymbol{\Sigma}^* \mathbf{f}_t^0 \right\|_2 \\ & \leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{M}_{\hat{\Lambda}}\|_2 \left\| \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \right\|_2 \|\hat{\Lambda}\|_2 \|\boldsymbol{\Sigma}^*\|_2 \sum_{\tau=1}^T \sum_{t=1}^T \|\mathbf{Z}_t\|_2 \|\mathbf{Z}_\tau\|_2 \|\boldsymbol{\varepsilon}_\tau\|_2 \|\mathbf{f}_t^0\|_2 \\ & \leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{M}_{\hat{\Lambda}}\|_2 \left\| \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \right\|_2 \|\hat{\Lambda}\|_2 \|\boldsymbol{\Sigma}^*\|_2 \left(\sum_{\tau=1}^T \|\mathbf{Z}_\tau\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{t=1}^T \|\mathbf{Z}_t\|_2^2 \right)^{\frac{1}{2}} \|\boldsymbol{\varepsilon}\|_F \|\mathbf{F}^0\|_F \\ & = O_P(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2), \end{aligned} \quad (\text{G.148})$$

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

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

Now,

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

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

$$\begin{aligned}
\mathcal{B}_{4.1.9} &= \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \left(-\frac{1}{nT} \sum_{\tau=1}^T \Lambda^0 \mathbf{f}_\tau^0 \left(\sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \Lambda^0 \mathbf{f}_\tau^0 \right)' \hat{\Lambda} \Sigma^* \right) \mathbf{f}_t^0 \\
&= \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \Lambda^0 \mathbf{F}^{0'} \mathbf{F}^0 \left(\sum_{q=1}^Q (\hat{\rho}_q - \rho_q^0) \mathbf{G}_q \Lambda^0 \right)' \hat{\Lambda} \Sigma^* \mathbf{f}_t^0 \\
&= \left(\frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} (\Lambda^0 - \hat{\Lambda} \mathbf{H}^{*-1}) \mathbf{F}^{0'} \mathbf{F}^0 \Lambda^{0'} (\Sigma^* \mathbf{f}_t^{0'} \otimes \mathbf{I}_n)' (\text{vec}(\mathbf{G}'_1 \hat{\Lambda}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\Lambda})) \right) \\
&\quad \times (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0). \tag{G.151}
\end{aligned}$$

Note that

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

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

$$\begin{aligned}
\mathcal{B}_{4.1.10} &= \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \left(-\frac{1}{nT} \sum_{\tau=1}^T \Lambda^0 \mathbf{f}_\tau^0 \left(\sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \boldsymbol{\varepsilon}_\tau \right)' \hat{\Lambda} \Sigma^* \right) \mathbf{f}_t^0 \\
&= \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \Lambda^0 \mathbf{F}^{0'} \boldsymbol{\varepsilon}' \left(\sum_{q=1}^Q (\hat{\rho}_q - \rho_q^0) \mathbf{G}_q \right)' \hat{\Lambda} \Sigma^* \mathbf{f}_t^0
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} (\Lambda^0 - \hat{\Lambda} \mathbf{H}^{*-1}) \mathbf{F}^{0'} \boldsymbol{\varepsilon}' (\boldsymbol{\Sigma}^* \mathbf{f}_t^{0'} \otimes \mathbf{I}_n)' (\text{vec}(\mathbf{G}'_1 \hat{\Lambda}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\Lambda})) \\
&\quad \times (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0). \tag{G.153}
\end{aligned}$$

Now,

$$\begin{aligned}
&\left\| \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} (\Lambda^0 - \hat{\Lambda} \mathbf{H}^{*-1}) \mathbf{F}^{0'} \boldsymbol{\varepsilon}' (\boldsymbol{\Sigma}^* \mathbf{f}_t^{0'} \otimes \mathbf{I}_n)' (\text{vec}(\mathbf{G}'_1 \hat{\Lambda}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\Lambda})) \right\|_2 \\
&\leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{M}_{\hat{\Lambda}}\|_2 \|\Lambda^0 - \hat{\Lambda} \mathbf{H}^{*-1}\|_2 \|\boldsymbol{\varepsilon} \mathbf{F}^0\|_2 \|\boldsymbol{\Sigma}^*\|_2 \|(\text{vec}(\mathbf{G}'_1 \hat{\Lambda}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\Lambda}))\|_2 \\
&\quad \times \sum_{t=1}^T \|\mathbf{Z}_t\|_2 \|\mathbf{f}_t^{0'}\|_2 \\
&\leq \frac{1}{nT} \frac{1}{nT} \|\mathbf{M}_{\hat{\Lambda}}\|_2 \|\Lambda^0 - \hat{\Lambda} \mathbf{H}^{*-1}\|_2 \|\boldsymbol{\varepsilon} \mathbf{F}^0\|_2 \|\mathbf{F}^0\|_F \|\boldsymbol{\Sigma}^*\|_2 \|\boldsymbol{\varepsilon}\|_2 \|(\text{vec}(\mathbf{G}'_1 \hat{\Lambda}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\Lambda}))\|_2 \\
&\quad \times \left(\sum_{t=1}^T \|\mathbf{Z}_t\|_2^2 \right)^{\frac{1}{2}} \\
&= O_P \left(\frac{Q\sqrt{P} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{T}} \right) + O_P \left(\frac{\sqrt{QP}}{\min\{\sqrt{nT}, T\}} \right), \tag{G.154}
\end{aligned}$$

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

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

Now,

$$\begin{aligned}
&\left\| \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \sum_{\tau=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_\tau (\hat{\Lambda} \boldsymbol{\Sigma}^* \mathbf{f}_t^0)' \mathbf{Z}_\tau \right\|_2 \\
&\leq \frac{1}{nT} \frac{1}{nT} \|\hat{\Lambda}\|_2 \|\boldsymbol{\Sigma}^*\|_2 \sum_{t=1}^T \sum_{\tau=1}^T \|\mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_\tau\|_2 \|\mathbf{f}_t^0\|_2 \|\mathbf{Z}_\tau\|_2 \\
&\leq \frac{1}{nT} \frac{1}{nT} \|\hat{\Lambda}\|_2 \|\boldsymbol{\Sigma}^*\|_2 \left(\sum_{t=1}^T \sum_{\tau=1}^T \|\mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_\tau\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{\tau=1}^T \|\mathbf{Z}_\tau\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{t=1}^T \|\mathbf{f}_t^0\|_2^2 \right)^{\frac{1}{2}} \\
&\leq \frac{1}{nT} \frac{1}{nT} \|\hat{\Lambda}\|_2 \|\boldsymbol{\Sigma}^*\|_2 \|\mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}\|_2 \left(\sum_{p=1}^P \|\mathbf{z}_p\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{\tau=1}^T \|\mathbf{Z}_\tau\|_2^2 \right)^{\frac{1}{2}} \left(\sum_{t=1}^T \|\mathbf{f}_t^0\|_2^2 \right)^{\frac{1}{2}}
\end{aligned}$$

$$= O_P \left(\frac{P}{\sqrt{\min\{n, T\}}} \right), \quad (\text{G.156})$$

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

$$\begin{aligned} \mathcal{B}_{4.1.12} &= \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \left(-\frac{1}{nT} \sum_{\tau=1}^T \varepsilon_\tau (\boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0)' \hat{\Lambda} \boldsymbol{\Sigma}^* \right) \mathbf{f}_t^0 \\ &= -\frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \sum_{\tau=1}^T \varepsilon_\tau (\boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0)' \hat{\Lambda} \boldsymbol{\Sigma}^* \mathbf{f}_t^0 \\ &= -\frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \varepsilon_\tau \mathbf{f}_\tau^{0'} \left(\frac{1}{nT} \mathbf{F}^{0'} \mathbf{F}^0 \right)^{-1} \mathbf{f}_t^0 \\ &= -\frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \varepsilon_\tau \varpi_{\tau t}^0 \end{aligned} \quad (\text{G.157})$$

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

$$\begin{aligned} \mathcal{B}_{4.1.14} &= \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \left(-\frac{1}{nT} \sum_{\tau=1}^T \varepsilon_\tau \left(\sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0 \right)' \hat{\Lambda} \boldsymbol{\Sigma}^* \right) \mathbf{f}_t^0 \\ &= -\frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon} \mathbf{F}^0 \boldsymbol{\Lambda}^{0'} \left(\sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \right)' \hat{\Lambda} \boldsymbol{\Sigma}^* \mathbf{f}_t^0 \\ &= \left(\frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon} \mathbf{F}^0 \boldsymbol{\Lambda}^{0'} (\boldsymbol{\Sigma}^* \mathbf{f}_t^{0'} \otimes \mathbf{I}_n)' (\text{vec}(\mathbf{G}'_1 \hat{\Lambda}), \dots, \text{vec}(\mathbf{G}'_Q \hat{\Lambda})) \right) (\boldsymbol{\rho}^0 - \hat{\boldsymbol{\rho}}). \end{aligned} \quad (\text{G.158})$$

Then,

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

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

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

$$\begin{aligned} \mathcal{B}_{4.1.17} &= \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \left(-\frac{1}{nT} \sum_{\tau=1}^T \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0 (\boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0)' \hat{\Lambda} \boldsymbol{\Sigma}^* \right) \mathbf{f}_t^0 \\ &= \frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \sum_{\tau=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \sum_{q=1}^Q (\hat{\rho}_q - \rho_q^0) \mathbf{G}_q \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0 (\boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0)' \hat{\Lambda} \boldsymbol{\Sigma}^* \mathbf{f}_t^0 \\ &= \frac{1}{nT} \frac{1}{nT} \left(\sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0 \right) (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0). \end{aligned} \quad (\text{G.160})$$

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

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

Now,

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

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 $\|(\text{vec}(\mathbf{G}_1\boldsymbol{\Lambda}^0), \dots, \text{vec}(\mathbf{G}_Q\boldsymbol{\Lambda}^0))\|_2 = O(\sqrt{Qn})$. Thus, $\mathcal{B}_{4.1.18} = O_P\left(\sqrt{\frac{QP}{T}}\right)(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0)$. Similar steps can be used to establish that $\mathcal{B}_{4.1.19}$, $\mathcal{B}_{4.1.20}$ and $\mathcal{B}_{4.1.21}$ are $O_P\left(Q\sqrt{P}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2\right)(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0)$, $O_P\left(\frac{Q\sqrt{P}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\sqrt{T}}\right)(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0)$ and $O_P\left(\sqrt{QP}\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2\right)(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0)$, respectively. Now consider term $\mathcal{B}_{4.1.22}$.

$$\begin{aligned}
\mathcal{B}_{4.1.22} &= \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\boldsymbol{\Lambda}}} \left(-\frac{1}{nT} \sum_{\tau=1}^T \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \boldsymbol{\varepsilon}_\tau (\boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0)' \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \right) \mathbf{f}_t^0 \\
&= -\frac{1}{nT} \frac{1}{nT} \sum_{t=1}^T \sum_{\tau=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\boldsymbol{\Lambda}}} \sum_{q=1}^Q (\rho_q^0 - \hat{\rho}_q) \mathbf{G}_q \boldsymbol{\varepsilon}_\tau (\boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0)' \hat{\boldsymbol{\Lambda}} \boldsymbol{\Sigma}^* \mathbf{f}_t^0 \\
&= \frac{1}{nT} \frac{1}{nT} \left(\sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\boldsymbol{\Lambda}}} \mathbf{G}_1 \boldsymbol{\varepsilon}_\tau, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\boldsymbol{\Lambda}}} \mathbf{G}_Q \boldsymbol{\varepsilon}_\tau \right) (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0).
\end{aligned} \tag{G.162}$$

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

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

Now,

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

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

$$\begin{aligned}
\mathbf{B}_{4.1} &= \left(\frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{Z}_\tau \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) - \frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \varepsilon_\tau \varpi_{\tau t}^0 \\
&\quad + \frac{1}{nT} \frac{1}{nT} \left(\sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0 \right) (\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 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \varepsilon_\tau, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \varepsilon_\tau \right) (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0) \\
&\quad + \mathbf{B}_{4.1.8} + \mathbf{B}_{4.1.13} + \mathbf{O}_P \left(\sqrt{QP} \|\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 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{Z}_\tau \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) - \frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \varepsilon_\tau \varpi_{\tau t}^0 \\
&\quad + \frac{1}{nT} \frac{1}{nT} \left(\sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0 \right) (\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 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \varepsilon_\tau, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \varepsilon_\tau \right) (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0) \\
&\quad + \boldsymbol{\Delta}_a + \mathbf{O}_P \left(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 + \frac{P}{\sqrt{\min\{n, T\}}} \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0), \tag{G.165}
\end{aligned}$$

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

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

Terms $\mathbf{B}_{4.3.1}$ and $\mathbf{B}_{4.3.3}$ can be written more compactly as $\mathbf{B}_{4.3.1} = \frac{1}{nT} \sum_{t=1}^T (\mathcal{M}_t^1)' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\Lambda}^0 \mathbf{f}_t^0$ and $\mathbf{B}_{4.3.3} = \frac{1}{nT} \sum_{t=1}^T (\mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\Lambda}^0 \mathbf{f}_t^0$ where $\mathcal{M}_t^1 := (\mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{f}_t^0, \dots, \mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{f}_t^0, \mathbf{0}_{n \times K})$ and

$\mathcal{M}_t^2 := (\mathbf{G}_1 \boldsymbol{\varepsilon}_t, \dots, \mathbf{G}_Q \boldsymbol{\varepsilon}_t, \mathbf{0}_{n \times K})$. Now note that

$$\begin{aligned} \left(\sum_{t=1}^T \|\mathcal{M}_t^1\|_2^2 \right)^{\frac{1}{2}} &\leq \left(\sum_{t=1}^T \|\mathcal{M}_t^1\|_F^2 \right)^{\frac{1}{2}} = \left(\sum_{q=1}^Q \|\mathbf{G}_q \boldsymbol{\Lambda}^0 \mathbf{F}^{0'}\|_F^2 \right)^{\frac{1}{2}} \\ &\leq \|\boldsymbol{\Lambda}^0 \mathbf{F}^{0'}\|_F^2 \sqrt{Q} \sqrt{\max_{1 \leq q \leq Q} \|\mathbf{G}_q\|_2} = O_P(\sqrt{QnT}), \end{aligned} \quad (\text{G.167})$$

and

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

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

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

and

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

Following analogous steps as those for terms $\mathcal{B}_{4.1.1}, \dots, \mathcal{B}_{4.1.6}, \mathcal{B}_{4.1.8}, \dots, \mathcal{B}_{4.2.5}$ yields the expression (the counterpart to (G.165))

$$\begin{aligned} &\mathcal{B}_{4.3.1} + \mathcal{B}_{4.3.3} \\ &= \left(\frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{Z}_{\tau} \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) - \frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_{\tau} \\ &\quad + \frac{1}{nT} \frac{1}{nT} \left(\sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{f}_{\tau}^0, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{f}_{\tau}^0 \right) \\ &\quad \times (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0) \\ &\quad + \frac{1}{nT} \frac{1}{nT} \left(\sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\varepsilon}_{\tau}, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\varepsilon}_{\tau} \right) \end{aligned}$$

$$\times(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0) + \boldsymbol{\Delta}_b + \mathbf{O}_P \left(\sqrt{Q}P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 + \frac{P}{\sqrt{\min\{n, T\}}} \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0), \quad (\text{G.169})$$

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

$$\begin{aligned} \mathcal{B}_4 &= \left(\frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{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 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_\tau \\ &+ \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)' \mathbf{M}_{\hat{\Lambda}} \mathbf{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)' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_\tau \\ &+ \frac{1}{nT} \frac{1}{nT} \left(\sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{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 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\varepsilon}_\tau, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \mathbf{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 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{f}_\tau^0, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{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 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_1 \boldsymbol{\varepsilon}_\tau, \dots, \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \mathbf{G}_Q \boldsymbol{\varepsilon}_\tau \right) \\ &\times (\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}^0) + \frac{1}{nT} \sum_{t=1}^T \mathbf{Z}'_t \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_t + \frac{1}{nT} \sum_{t=1}^T (\mathcal{M}_t^1 + \mathcal{M}_t^2)' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_t \\ &+ \boldsymbol{\Delta}_a + \boldsymbol{\Delta}_b + \mathbf{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} \quad (\text{G.170})$$

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

$$\begin{aligned} \mathcal{B}_4 &= \frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t^*)' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_t - \frac{1}{nT} \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T (\mathbf{Z}_t^*)' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon}_\tau \varpi_{\tau t}^0 \\ &+ \frac{1}{nT} \frac{1}{T} \left(\begin{array}{ccc} \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathbf{W}_1 \mathbf{y}_t)' \mathbf{M}_{\hat{\Lambda}} \mathbf{W}_1 \mathbf{y}_\tau & \cdots & \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 (\mathbf{W}_1 \mathbf{y}_t)' \mathbf{M}_{\hat{\Lambda}} \mathbf{x}_{K\tau} \\ \vdots & \ddots & \vdots \\ \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{x}'_{Kt} \mathbf{M}_{\hat{\Lambda}} \mathbf{W}_1 \mathbf{y}_\tau & \cdots & \sum_{t=1}^T \sum_{\tau=1}^T \varpi_{\tau t}^0 \mathbf{x}'_{Kt} \mathbf{M}_{\hat{\Lambda}} \mathbf{x}_{K\tau} \end{array} \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \end{aligned}$$

$$\begin{aligned}
& + \mathcal{O}_P \left(\sqrt{Q}P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 + \frac{P}{\sqrt{\min\{n, T\}}} \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + \boldsymbol{\Delta}_1 \\
& = \frac{1}{nT} \boldsymbol{Z}^{*\prime} (\boldsymbol{M}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}) \text{vec}(\boldsymbol{\varepsilon}) + \frac{1}{nT} \boldsymbol{Z}^{*\prime} (\boldsymbol{P}_{\boldsymbol{F}^0} \otimes \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}) \boldsymbol{Z}^* (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
& + \mathcal{O}_P \left(\sqrt{Q}P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2 + \frac{P}{\sqrt{\min\{n, T\}}} \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + \boldsymbol{\Delta}_1. \tag{G.171}
\end{aligned}$$

This completes the proof. \square

Proof of Lemma F.3(v). For brevity, let $\hat{\sigma}^2 := \hat{\sigma}^2(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}})$, and recall also that $\hat{\sigma}^2(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Lambda}}) = \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{S}(\hat{\boldsymbol{\rho}}) \boldsymbol{y}_t - \boldsymbol{X}_t \hat{\boldsymbol{\beta}})' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} (\boldsymbol{S}(\hat{\boldsymbol{\rho}}) \boldsymbol{y}_t - \boldsymbol{X}_t \hat{\boldsymbol{\beta}})$. Using Lemma A.2(i) and the true DGP, one obtains $\boldsymbol{S}(\hat{\boldsymbol{\rho}}) \boldsymbol{y}_t - \boldsymbol{X}_t \hat{\boldsymbol{\beta}} = \boldsymbol{Z}_t^* (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \boldsymbol{\Lambda}^0 \boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t$, where $\boldsymbol{Z}_t^* := (\boldsymbol{W}_1 \boldsymbol{y}_t, \dots, \boldsymbol{W}_q \boldsymbol{y}_t, \boldsymbol{X}_t)$. Thus the expression for $\hat{\sigma}^2$ can be expanded to give

$$\begin{aligned}
\hat{\sigma}^2 & = (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \left(\frac{1}{nT} \sum_{t=1}^T (\boldsymbol{Z}_t^*)' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{Z}_t^* (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) + \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{Z}_t^*)' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 \boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) \right) \\
& + (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \left(\frac{1}{nT} \sum_{t=1}^T (\boldsymbol{Z}_t^*)' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 \boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) \right) + \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{\Lambda}^0 \boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t)' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 \boldsymbol{f}_t^0 + \boldsymbol{\varepsilon}_t) \\
& =: (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \boldsymbol{l}_1 + (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \boldsymbol{l}_2 + \boldsymbol{l}_3. \tag{G.172}
\end{aligned}$$

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

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

where $\boldsymbol{Z}^* := (\boldsymbol{Z}_1^{*\prime}, \dots, \boldsymbol{Z}_T^{*\prime})'$. An analogous expansion of term \boldsymbol{l}_3 in equation (G.172) yields

$$\begin{aligned}
\boldsymbol{l}_3 & = \text{tr}((\boldsymbol{\Lambda}^0 \boldsymbol{F}^{0'} + \boldsymbol{\varepsilon})' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) + \boldsymbol{\mathcal{H}}^*(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
& + \boldsymbol{\Delta}_2 + \left(\mathcal{O}_P(\sqrt{Q}P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathcal{O}_P \left(\frac{P}{\sqrt{\min\{n, T\}}} \right) \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
& = \text{tr}(\boldsymbol{\varepsilon}' \boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}} \boldsymbol{\varepsilon} \boldsymbol{M}_{\boldsymbol{F}^0}) + \boldsymbol{\mathcal{H}}^*(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \\
& + \boldsymbol{\Delta}_2 + \left(\mathcal{O}_P(\sqrt{Q}P \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathcal{O}_P \left(\frac{P}{\sqrt{\min\{n, T\}}} \right) \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0), \tag{G.174}
\end{aligned}$$

since $\text{tr}((\mathbf{\Lambda}^0 \mathbf{F}^{0'})' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) = 0$, and where $\mathcal{H}^* := \frac{1}{nT} \text{vec}(\mathbf{\Lambda}^0 \mathbf{F}^{0'} + \boldsymbol{\varepsilon})' (\mathbf{P}_{\mathbf{F}^0} \otimes \mathbf{M}_{\hat{\Lambda}}) \mathcal{Z}$ and Δ_2 is a term of the same order as Δ_1 . Combining (G.172), (G.173) and (G.174),

$$\begin{aligned} \hat{\sigma}^2 &= (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}})' \left(\frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t^*)' \mathbf{M}_{\hat{\Lambda}} \mathbf{Z}_t^* - \mathcal{H} \right) (\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}) \\ &\quad + \frac{2}{nT} \mathcal{Z}^{*'} (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\hat{\Lambda}}) \text{vec}(\boldsymbol{\varepsilon}) + \text{tr}(\boldsymbol{\varepsilon}' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) + \Delta_2 \\ &\quad + \left(\mathcal{O}_P(\sqrt{QP} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2) + \mathcal{O}_P \left(\frac{P}{\sqrt{\min\{n, T\}}} \right) \right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0), \end{aligned} \quad (\text{G.175})$$

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

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

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

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

$$\begin{aligned}
& + \frac{1}{nT} \begin{pmatrix} \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \vdots \\ \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \mathbf{0}_{K \times 1} \end{pmatrix} + \mathcal{O}_P\left(\frac{1}{\sqrt{nT}}\right) \\
& =: \frac{1}{nT} \mathbf{Z}'(\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\Lambda^0}) \text{vec}(\boldsymbol{\varepsilon}) + \frac{1}{nT} \begin{pmatrix} \text{tr}((\mathbf{G}_1 \boldsymbol{\Lambda}^0 \mathbf{F}^{0'})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \vdots \\ \text{tr}((\mathbf{G}_Q \boldsymbol{\Lambda}^0 \mathbf{F}^{0'})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \mathbf{0}_{K \times 1} \end{pmatrix} \\
& + \frac{1}{nT} \begin{pmatrix} \text{tr}((\mathbf{G}_1 \boldsymbol{\varepsilon})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \vdots \\ \text{tr}((\mathbf{G}_Q \boldsymbol{\varepsilon})' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\ \mathbf{0}_{K \times 1} \end{pmatrix} + \mathcal{O}_P\left(\frac{1}{\sqrt{nT}}\right), \tag{G.177}
\end{aligned}$$

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

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

Lastly, using Lemma F.3(vi)

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

and

$$\begin{aligned}
& \frac{1}{nT} \text{tr}(\boldsymbol{\varepsilon}' \mathbf{M}_{\hat{\Lambda}} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) \\
& = \frac{1}{nT} \text{tr}(\boldsymbol{\varepsilon}' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0}) + \frac{1}{nT} \text{tr}(\boldsymbol{\varepsilon}' (\mathbf{P}_{\Lambda^0} - \mathbf{P}_{\hat{\Lambda}}) \boldsymbol{\varepsilon} \mathbf{M}_{\mathbf{F}^0})
\end{aligned}$$

$$= \frac{1}{nT} \text{tr}(\boldsymbol{\varepsilon}' \mathbf{M}_{\Lambda^0} \boldsymbol{\varepsilon} \mathbf{M}_{F^0}) + O_P \left(\frac{\sqrt{Q} \|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2}{\min\{n, T\}} \right) + O_P \left(\frac{1}{\min\{n^{1.5}, T^{1.5}\}} \right), \quad (\text{G.180})$$

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

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

□

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

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

Now,

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

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

$$\mathbb{E} [\|\mathbf{l}_{1.1}\|_2^2] = \mathbb{E} \left[\sum_{q=1}^Q \left(\sum_{t=1}^T \boldsymbol{\varepsilon}_t' \mathbf{G}_q' \boldsymbol{\varepsilon}_t - \sigma_0^2 \text{tr}(\mathbf{G}_q) \right)^2 \right]$$

$$\begin{aligned}
&= \sum_{q=1}^Q T \left((\mathcal{M}_4^\varepsilon - 3\sigma_0^4) \sum_{i=1}^n (G_q)_{ii}^2 + \sigma_0^4 (\text{tr}(\mathbf{G}_q \mathbf{G}_q) + \text{tr}(\mathbf{G}'_q \mathbf{G}_q)) \right) \\
&= O(QnT), \tag{G.183}
\end{aligned}$$

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

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

where Lemma F.1(ii) has been applied, with Λ^0 replaced by $\mathbf{G}'_q \Lambda^0$, to establish that $\sum_{q=1}^Q \|(\mathbf{G}_q \boldsymbol{\varepsilon})' \Lambda^0\|_2^2 = O_P(QnT)$. Thus $\|\mathbf{l}_2\|_2 = O_P(\sqrt{QT})$. Next for term \mathbf{l}_3 ,

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

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

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

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

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

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

Proof of Lemma F.4(i).

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

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

Proof of Lemma F.4(ii).

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

$$\begin{aligned}
&\leq \|\boldsymbol{\beta}^0 - \hat{\boldsymbol{\beta}}\|_2^2 Q \max_{1 \leq q \leq Q} \{\|\mathbf{G}_q\|_2^2\} \sum_{k=1}^K \|\boldsymbol{\mathcal{X}}_k\|_F^2 \\
&= O_P(\|\boldsymbol{\theta}^0 - \hat{\boldsymbol{\theta}}\|_2^2) O_P(QKnT), \tag{G.189}
\end{aligned}$$

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

Proof of Lemma F.4(iii).

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

using Lemmas A.2(vii) and A.2(viii). \square

Proof of Lemma F.4(iv).

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

$$= O_P(\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2^2) O_P(QKnT), \quad (\text{G.191})$$

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

Proof of Lemma F.4(v).

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

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

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 \mathbf{F} , and with σ^2 concentrated out, is denoted $\mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \mathbf{F})$. Dropping the constant this is given by

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

Substituting the true DGP $\mathbf{y}_t = \mathbf{S}^{-1}(\mathbf{X}_t \boldsymbol{\beta}^0 + \boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \boldsymbol{\varepsilon}_t)$ into (H.1) and applying Lemma A.2(i) results in

$$\begin{aligned} \mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \mathbf{F}) &= \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) \\ &\quad - \frac{1}{2} \log \left(\mathbb{E} \left[\frac{1}{nT} \sum_{t=1}^T (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) + \mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\Lambda}^0 \mathbf{f}_t^0 + \mathbf{S}(\boldsymbol{\rho}) \mathbf{S}^{-1} \boldsymbol{\varepsilon}_t - \boldsymbol{\Lambda} \mathbf{f}_t)' \right] \right) \end{aligned}$$

$$\times (\mathbf{Z}_t(\boldsymbol{\theta}^0 - \boldsymbol{\theta}) + \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\Lambda}^0\mathbf{f}_t^0 + \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon}_t - \boldsymbol{\Lambda}\mathbf{f}_t)]. \quad (\text{H.2})$$

To begin, it is shown that for any $(\boldsymbol{\theta}, \boldsymbol{\Lambda}\mathbf{F}) \neq (\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0\mathbf{F}^0)$, $\mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \mathbf{F}) < \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0)$. First of all,

$$\begin{aligned} \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0) &= \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log \left(\mathbb{E} \left[\frac{1}{nT} \sum_{t=1}^T \boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t \right] \right) \\ &= \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2), \end{aligned} \quad (\text{H.3})$$

where the second line follows by Assumption 1.1. Next, using Assumption 1, and concentrating out \mathbf{F} and $\boldsymbol{\Lambda}^0$, gives the inequality

$$\begin{aligned} \mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \mathbf{F}) &\leq \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) - \frac{1}{2} \log \left(\frac{1}{nT} \mathbb{E} [\text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon})] \right. \\ &\quad \left. + \mathbb{E} \left[\text{tr} \left(\mathbf{M}_{\mathbf{F}^0} \left(\sum_{p=1}^P (\theta_p^0 - \theta_p) \boldsymbol{z}_p \right)' \mathbf{M}_{\boldsymbol{\Lambda}} \left(\sum_{p=1}^P (\theta_p^0 - \theta_p) \boldsymbol{z}_p \right) \right) \right] \right). \end{aligned} \quad (\text{H.4})$$

By Lemma 9 in Yu et al. (2008), $\frac{1}{nT} \mathbb{E} [\text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}\boldsymbol{\varepsilon})] = \frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})$. Applying this, and then rearranging (H.4), gives

$$\begin{aligned} \mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \mathbf{F}) &\leq \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) \\ &\quad - \frac{1}{2} \log \left(\frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) + \mathbb{E} \left[\frac{1}{nT} \boldsymbol{Z}' (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\boldsymbol{\Lambda}}) \boldsymbol{Z} \right] \right). \end{aligned} \quad (\text{H.5})$$

For simplicity denote $\mathbb{E} \left[\frac{1}{nT} \boldsymbol{Z}' (\mathbf{M}_{\mathbf{F}^0} \otimes \mathbf{M}_{\boldsymbol{\Lambda}}) \boldsymbol{Z} \right]$ by \mathbf{M} . Now, $\boldsymbol{\theta}^0$ is a unique global maximiser of the unpenalised expected likelihood for any $\boldsymbol{\Lambda}, \mathbf{F}$, if for any $\boldsymbol{\theta} \neq \boldsymbol{\theta}^0$, $\mathbb{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}, \mathbf{F}) < \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}, \mathbf{F})$. Using (H.3) and (H.5), this inequality holds when

$$\begin{aligned} \frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) - \frac{1}{2} \log \left(\frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) + (\boldsymbol{\theta}^0 - \boldsymbol{\theta})' \mathbf{M} (\boldsymbol{\theta}^0 - \boldsymbol{\theta}) \right) \\ < \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2). \end{aligned} \quad (\text{H.6})$$

Note that $-\frac{1}{n} \log(\det(\mathbf{S}(\boldsymbol{\rho}))) + \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2) = -\frac{1}{2} \log(\sigma_0^2 \det((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})^{\frac{1}{n}})$. Therefore (H.6) becomes

$$\begin{aligned} -\frac{1}{2} \log \left(\frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) + (\boldsymbol{\theta}^0 - \boldsymbol{\theta})' \mathbf{M} (\boldsymbol{\theta}^0 - \boldsymbol{\theta}) \right) \\ < -\frac{1}{2} \log(\sigma_0^2 \det((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})' \mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})^{\frac{1}{n}}). \end{aligned} \quad (\text{H.7})$$

Multiplying by $-\frac{1}{2}$ and exponentiating yields the condition

$$\frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) + (\boldsymbol{\theta}^0 - \boldsymbol{\theta})'\mathbf{M}(\boldsymbol{\theta}^0 - \boldsymbol{\theta})) > \sigma_0^2 \det((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})^{\frac{1}{n}}. \quad (\text{H.8})$$

The matrix \mathbf{M} is positive definite by Assumption ID.2 and, moreover, by Lemma A.1, $\frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) \geq \sigma_0^2 \det((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})^{\frac{1}{n}}$. Hence,

$$\frac{\sigma_0^2}{n} \text{tr}((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1}) - \sigma_0^2 \det((\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})'\mathbf{S}(\boldsymbol{\rho})\mathbf{S}^{-1})^{\frac{1}{n}} + (\boldsymbol{\theta}^0 - \boldsymbol{\theta})'\mathbf{M}(\boldsymbol{\theta}^0 - \boldsymbol{\theta})) > 0, \quad (\text{H.9})$$

for any $\boldsymbol{\theta} \neq \boldsymbol{\theta}^0$ and the expected likelihood must be uniquely maximised at $\boldsymbol{\theta}^0$ for any $\boldsymbol{\Lambda}, \mathbf{F}$. Now, evaluated at $\boldsymbol{\theta}^0$ and with $\boldsymbol{\Lambda}\mathbf{F}' \neq \boldsymbol{\Lambda}^0\mathbf{F}^{0'}$,

$$\begin{aligned} \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}, \mathbf{F}) &\leq \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log \left(\mathbb{E} \left[\frac{1}{nT} \sum_{t=1}^T \varepsilon'_t \varepsilon_t + \frac{1}{nT} \sum_{t=1}^T (\boldsymbol{\Lambda}^0 \mathbf{f}_t^0 - \boldsymbol{\Lambda} \mathbf{f}_t)' (\boldsymbol{\Lambda}^0 \mathbf{f}_t^0 - \boldsymbol{\Lambda} \mathbf{f}_t) \right] \right) \\ &= \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log \left(\sigma_0^2 + \mathbb{E} \left[\frac{1}{nT} \text{tr}((\boldsymbol{\Lambda}^0 \mathbf{F}^{0'} - \boldsymbol{\Lambda} \mathbf{F}')' (\boldsymbol{\Lambda}^0 \mathbf{F}^{0'} - \boldsymbol{\Lambda} \mathbf{F}')) \right] \right). \end{aligned} \quad (\text{H.10})$$

The trace term in (H.10) can be equivalently written as $\text{vec}(\boldsymbol{\Lambda}^0 \mathbf{F}^{0'} - \boldsymbol{\Lambda} \mathbf{F}')' (\mathbf{I}_T \otimes \mathbf{I}_n) \text{vec}(\boldsymbol{\Lambda}^0 \mathbf{F}^{0'} - \boldsymbol{\Lambda} \mathbf{F}')$. Since the matrix $(\mathbf{I}_T \otimes \mathbf{I}_n)$ is positive definite, this term is strictly positive as long as $\boldsymbol{\Lambda}^0 \mathbf{F}^{0'} \neq \boldsymbol{\Lambda} \mathbf{F}'$ and therefore

$$\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}, \mathbf{F}) < \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0) = \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \log(\sigma_0^2) \quad (\text{H.11})$$

and the expected likelihood is maximised where $\boldsymbol{\Lambda}^0 \mathbf{F}^{0'} = \boldsymbol{\Lambda} \mathbf{F}'$, which is sufficient for the identification of $\boldsymbol{\Lambda}^0 \mathbf{F}^{0'}$. Identification of σ_0^2 is then straightforward to show since, omitting the constant, $\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \sigma_0^2) = -\frac{1}{2} \log(\sigma_0^2) + \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2}$ and $\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \sigma^2) = -\frac{1}{2} \log(\sigma^2) + \frac{1}{n} \log(\det(\mathbf{S})) - \frac{1}{2} \frac{\sigma_0^2}{\sigma^2}$. Thus $\mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \sigma^2) < \mathbb{L}(\boldsymbol{\theta}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \sigma_0^2)$ holds if

$$-\frac{1}{2} \log(\sigma^2) - \frac{1}{2} \frac{\sigma_0^2}{\sigma^2} < -\frac{1}{2} \log(\sigma_0^2) - \frac{1}{2} \quad (\text{H.12})$$

or

$$\log \left(\frac{\sigma_0^2}{\sigma^2} \right) < \frac{\sigma_0^2}{\sigma^2} - 1. \quad (\text{H.13})$$

Using $\log(x) < x - 1$ for $x > 0$ 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 $\mathbf{S}(\boldsymbol{\rho})$: For simplicity, let $\mathbf{W}^* := \sum_{q=1}^Q \rho_q \mathbf{W}_q$. A sufficient condition for the invertibility of $\mathbf{S}(\boldsymbol{\rho}) = \mathbf{I}_n - \mathbf{W}^*$ is that $\|\mathbf{W}^*\| < 1$ for some norm $\|\cdot\|$.³ Since

$$\|\mathbf{W}^*\| \leq \sum_{q=1}^Q |\rho_q| \max_{1 \leq q \leq Q} \|\mathbf{W}_q\|, \quad (\text{I.1})$$

$\mathbf{S}(\boldsymbol{\rho})$ will be invertible if $\sum_{q=1}^Q |\rho_q| < (\max_{1 \leq q \leq Q} \|\mathbf{W}_q\|)^{-1}$. In particular, since all the \mathbf{W}_q are row normalised (so that $\|\mathbf{W}_q\|_\infty = 1$), this condition reduces to $\sum_{q=1}^Q |\rho_q| < 1$, which is satisfied in the MC design.

$\mathbf{S}(\boldsymbol{\rho})$ is UB: Since $\|\mathbf{S}(\boldsymbol{\rho})\| = \|\mathbf{I}_n - \mathbf{W}^*\| \leq \|\mathbf{I}_n\| + \|\mathbf{W}^*\|$, $\mathbf{S}(\boldsymbol{\rho})$ is UB if \mathbf{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.

$\mathbf{S}^{-1}(\boldsymbol{\rho})$ is UB: If $\|\mathbf{W}^*\| < 1$ for some norm $\|\cdot\|$, then $\mathbf{S}^{-1}(\boldsymbol{\rho}) = \sum_{h=0}^{\infty} (\mathbf{W}^*)^h$ and therefore

$$\|\mathbf{S}^{-1}(\boldsymbol{\rho})\| \leq \sum_{h=0}^{\infty} \|(\mathbf{W}^*)^h\| \leq \sum_{h=0}^{\infty} \|\mathbf{W}^*\|^h. \quad (\text{I.2})$$

Under the condition $\sum_{q=1}^Q |\rho_q| < 1$, $\|\mathbf{W}^*\|_\infty < 1$ (see above) and therefore $\|\mathbf{S}^{-1}(\boldsymbol{\rho})\|_\infty$ 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), $\|\mathbf{W}^*\|_1$ 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 $\mathbf{S}^{-1}(\boldsymbol{\rho})$ 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.

3.2, 5, 6.6 In simulations, γ_ρ and γ_β are chosen using the information criterion described in Section 4.1. The parameter ζ is set equal to 4.

³Easily verifiable by considering the Neumann series of $(\mathbf{I}_n - \mathbf{W})^{-1}$.

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 $R \geq R^0$.

4.4 In the design, the number of parameters adheres to a sequence, indexed say $j = 1, 2, \dots$, 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, $\lim_{j \rightarrow \infty} \frac{P(j)}{\min\{n(j), T(j)\}} = \lim_{j \rightarrow \infty} \frac{8+4j}{25 \times 2^{j-1}} = 0$.

6.1, 8.1 Considering the same sequence as in 4.4, note that $\lim_{j \rightarrow \infty} \frac{P^5(j)}{\min\{n(j), T(j)\}} = \lim_{j \rightarrow \infty} \frac{(8+4j)^5}{25 \times 2^{j-1}} = 0$, whereby Assumptions 6.1 and 8.1 are satisfied.

8.2 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 the previous point, the sequences $\min\{n^{1/4}, T^{1/4}\}/\sqrt{PQ} \rightarrow \infty$ and $P/\min\{n^{1/4}, T^{1/4}\} \rightarrow 0$.

J Additional Tables

Lemma 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	25	0.0131	0.0009	-	0.0583	0.0729	-	-0.0899	-0.0286	-	-0.0013	0.0012
	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
50	25	0.0122	-0.0011	-0.0010	0.0635	0.0790	-	-0.0905	-0.0411	-	-0.0007	0.0008
	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
100	25	0.0115	-0.0022	0.0007	0.0671	0.0774	-	-0.0830	-0.0384	-0.0432	-0.0012	0.0012
	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	25	0.631	0.885	-	0.465	0.337	-	0.410	0.817	-	0.914	0.908
	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
50	25	0.457	0.907	0.894	0.159	0.073	-	0.172	0.588	-	0.902	0.904
	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
100	25	0.325	0.865	0.921	0.021	0.009	-	0.109	0.493	0.408	0.880	0.880
	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 ($R = 6$)

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	-0.0003	-	0.0006	-0.0010	-	-0.0021	0.0036	-	-0.0003	-0.0001
	50	0	-0.0003	-	0.0001	-0.0006	-	-0.0023	0.0030	-	-0.0002	0.0002
	100	-0.0001	-0.0001	-	0.0003	-0.0010	0.0008	-0.0002	0.0009	-	-0.0002	0.0002
50	25	0.0001	-0.0002	0	0.0010	-0.0004	-	-0.0010	0.0006	-	-0.0001	0.0002
	50	0.0001	-0.0004	0.0001	0.0002	-0.0012	-	-0.0005	0.0017	-	-0.0002	0.0002
	100	0.0001	-0.0002	0	0.0004	-0.0005	0.0005	-0.0003	0.0005	-	-0.0001	0.0001
100	25	0.0001	-0.0001	0	-0.0003	-0.0010	-	-0.0003	0.0015	-0.0001	-0.0003	0.0003
	50	0	-0.0001	0	0.0006	-0.0006	-	-0.0002	0.0016	-0.0007	-0.0002	0.0002
	100	0.0001	-0.0001	0	0.0001	-0.0004	0.0003	0.0001	0.0005	-0.0003	-0.0001	0.0001

Table 10: Coverage of nonzero parameter estimates ($R = 6$)

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	0.735	0.750	-	0.742	0.751	-	0.736	0.764	-	0.745	0.751
	50	0.796	0.788	-	0.830	0.821	-	0.819	0.805	-	0.826	0.830
	100	0.832	0.858	-	0.863	0.844	0.847	0.848	0.854	-	0.854	0.852
50	25	0.798	0.811	0.820	0.829	0.816	-	0.821	0.821	-	0.809	0.808
	50	0.858	0.855	0.861	0.888	0.873	-	0.850	0.876	-	0.863	0.860
	100	0.903	0.887	0.907	0.891	0.885	0.889	0.885	0.901	-	0.897	0.905
100	25	0.877	0.844	0.864	0.856	0.858	-	0.855	0.846	0.852	0.867	0.851
	50	0.913	0.899	0.904	0.870	0.903	-	0.896	0.893	0.878	0.895	0.900
	100	0.912	0.917	0.908	0.939	0.916	0.910	0.915	0.929	0.911	0.909	0.905

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	25	-0.0001	-0.0020	-	-0.0009	0.0008	-	0.0020	0.0066	-	-0.0001	0.0006
	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
50	25	0.0003	-0.0005	-0.0001	0.0014	-0.0012	-	-0.0003	0.0015	-	-0.0004	0.0005
	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
100	25	-0.0001	-0.0004	0.0001	0.0005	-0.0010	-	0.0002	0.0030	-0.0011	-0.0003	0.0003
	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	25	0.462	0.450	-	0.484	0.471	-	0.469	0.464	-	0.505	0.484
	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
50	25	0.621	0.629	0.655	0.621	0.661	-	0.659	0.651	-	0.659	0.658
	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
100	25	0.728	0.740	0.719	0.732	0.708	-	0.711	0.726	0.729	0.698	0.707
	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	25	99.9	-	90	-	90.4	-	100	-	-
	50	99.7	-	100	100	99.4	-	100	-	-
	100	99.7	-	100	100	100	-	99.7	-	-
50	25	100	-	100	-	99.8	99.8	100	100	-
	50	100	-	100	100	100	100	100	100	-
	100	99.6	-	100	100	100	100	99.6	99.6	-
100	25	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	25	0.1	0.2	-	0	0	-	0	0	-	0	0
	50	0.1	0.3	-	0	0	-	0	0	-	0	0
	100	0	0	-	0	0	0	0	0	-	0	0
50	25	0	0	0	0	0	-	0	0	-	0	0
	50	0	0	0	0	0	-	0	0	-	0	0
	100	0	0	0	0	0	0	0	0	-	0	0
100	25	0	0	0	0	0	-	0	0	0	0	0
	50	0	0	0	0	0	-	0	0	0	0	0
	100	0	0	0	0	0	0	0	0	0	0	0

Table 15: True number of factors is selected % ($R = 10$)

T	25			50			100		
n	IC1	IC2	IC3	IC1	IC2	IC3	IC1	IC2	IC3
25	0	96.6	86	23.5	99.8	99.8	100	100	100
50	37.2	100	98.4	1.9	100	100	100	100	100
100	100	100	100	100	100	100	99.2	100	100

Lemma J.2 Gamma(0.5,1)

Table 16: 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	25	0.0006	-0.0007	-	0.0009	-0.0006	-	-0.0027	0.0023	-	-0.0003	0.0004
	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
50	25	0.0002	-0.0002	-0.0001	-0.0003	-0.0005	-	-0.0006	0.0015	-	-0.0003	0.0004
	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
100	25	0.0002	-0.0004	0.0002	0.0003	-0.0001	-	-0.0017	0.0028	-0.0014	-0.0003	0.0003
	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 17: 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	0.891	0.896	-	0.889	0.896	-	0.901	0.906	-	0.893	0.895
	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
50	25	0.936	0.931	0.925	0.922	0.930	-	0.918	0.919	-	0.928	0.930
	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
100	25	0.925	0.944	0.919	0.907	0.927	-	0.914	0.909	0.929	0.932	0.932
	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

Lemma 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	25	-0.0001	-0.0004	-	0.0013	-0.0021	-	-0.0012	0.0022	-	-0.0001	0.0001
	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
50	25	0.0003	-0.0003	0.0001	0	-0.0001	-	-0.0018	0.0017	-	-0.0002	0.0002
	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
100	25	0	0	0	0.0002	0.0003	-	-0.0005	0.0010	-0.0007	-0.0002	0.0002
	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	25	0.906	0.910	-	0.914	0.899	-	0.893	0.912	-	0.931	0.928
	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
50	25	0.898	0.912	0.918	0.924	0.925	-	0.913	0.940	-	0.924	0.921
	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
100	25	0.929	0.928	0.923	0.926	0.924	-	0.935	0.940	0.926	0.916	0.916
	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

Lemma J.4 Laplace(0,1)

Table 20: 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 21: 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

Lemma J.5 ChiSq(3)

Table 22: 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 23: 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	0.906	-	0.901	0.898	-	0.893	0.896	-	0.903	0.899
25	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
	25	0.927	0.913	0.910	0.914	0.911	-	0.917	0.923	-	0.917	0.922
50	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
	25	0.914	0.935	0.931	0.940	0.923	-	0.930	0.940	0.926	0.938	0.933
100	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

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.

Lemma 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 24: 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 25: 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 26: 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 26 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 decreases as sample size increases.

Lemma 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 27.

Table 27: 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 $x_{kit}^* = \nu + \sum_{r=1}^{R^0} \lambda_{ir}^0 f_{rt}^0 + e_{it}$ with ν being uniformly drawn from the integers $\{-10, \dots, 10\}$ 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 28–31 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 28: 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
3	25	-0.0006	-0.0008	-	0.0002	-0.0001	-0.0001	0.0005
	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
5	25	-0.0008	-0.0006	-	-0.0005	-0.0008	-0.0002	0.0005
	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
10	25	-0.0107	-0.0110	-	-0.0024	0.0012	-0.0004	0.0050
	50	-0.0002	-0.0005	-0.0004	0.0005	-0.0002	0	0.0009
	100	0	0	0	0	0	0	0.0002
15	25	-0.0287	-0.0297	-	-0.0033	0.0016	-0.0010	0.0144
	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 29: 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 30: 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 31: 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 borne

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.
- 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.
- Lewbel, A., 2019. The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature* 57 (4), 835–903.
- 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.
- Newey, W. K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R. F., McFadden, D. L. (Eds.), *Handbook of Econometrics*. Vol. 4. Elsevier, Ch. 36, pp. 2111–2245.
- 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.