Supplementary Material for

"Panel Data Models with Interactive Fixed Effects and Multiple Structural Breaks"

This supplemental document provides the proofs of all the technical lemmas in Appendix B of the main document.

C Proofs of the technical lemmas

In this appendix we give the detailed proofs of the technical lemmas used in Appendix B. Before proving Lemma B.1 on the convergence rates of $\dot{\beta}_t$, we give some preliminary results. Let $\boldsymbol{b} = (b_1', b_2', ..., b_T')'$ where b_t is a p-dimensional column vector and let C be a positive constant whose value may change from line to line. Recall that $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$.

Lemma C.1 *Suppose that Assumption 1 in Appendix A holds. Then we have*

(i) $\sup_{\boldsymbol{b}} \sup_{\boldsymbol{\Lambda}} \left| \frac{1}{NT} \sum_{t=1}^T b_t' X_t' \boldsymbol{M}_{\boldsymbol{\Lambda}} \varepsilon_t \right| = O_P(pN^{-1/2} + p^{1/2} T^{-1/2})$,

(ii) $\sup_{\boldsymbol{\Lambda}} \left| \frac{1}{NT} \sum_{t=1}^T f_t^{0\prime} \boldsymbol{\Lambda}^{0\prime} \boldsymbol{M}_{\boldsymbol{\Lambda}} \varepsilon_t \right| = O_P(\delta_{NT}^{-1})$,

(iii) $\sup_{\boldsymbol{\Lambda}} \left| \frac{1}{NT} \sum_{t=1}^T \varepsilon_t' \boldsymbol{P}_{\boldsymbol{\Lambda}} \varepsilon_t \right| = O_P(\delta_{NT}^{-2})$,

(iv) $\frac{1}{NT} \sum_{t=1}^T \varepsilon_t' \boldsymbol{P}_{\boldsymbol{\Lambda}^0} \varepsilon_t = O_P(N^{-1})$,

where $\sup_{\boldsymbol{b}}$ *is taken with respect to* $\boldsymbol{b}$ *such that* $\|\boldsymbol{b}\| \le C(pT)^{1/2}$ *and* $\sup_{\boldsymbol{\Lambda}}$ *is taken with respect to* $\boldsymbol{\Lambda}$ *such that* $\frac{1}{N}\boldsymbol{\Lambda}'\boldsymbol{\Lambda} = \boldsymbol{I}_{R_0}$.

Proof of Lemma C.1. **(i)** Note that $\frac{1}{NT} \sum_{t=1}^T b_t' X_t' \boldsymbol{M}_{\boldsymbol{\Lambda}} \varepsilon_t = \frac{1}{NT} \sum_{t=1}^T b_t' X_t' \varepsilon_t - \frac{1}{N^2 T} \sum_{t=1}^T b_t' X_t' \boldsymbol{\Lambda}\boldsymbol{\Lambda}' \varepsilon_t$ if $\frac{1}{N}\boldsymbol{\Lambda}'\boldsymbol{\Lambda} = \boldsymbol{I}_{R_0}$. By Assumption 1(iii) and the Cauchy-Schwarz inequality, we have

$$\Big| \sum_{t=1}^T b_t' X_t' \varepsilon_t \Big| = \Big(\sum_{t=1}^T \|b_t\|^2 \Big)^{1/2} \cdot \Big(\sum_{t=1}^T \|X_t'\varepsilon_t\|^2 \Big)^{1/2} = O_P\big(pTN^{1/2}\big) \tag{C.1}$$

for $\|\boldsymbol{b}\|^2 = \sum_{t=1}^T \|b_t\|^2 \le CpT$. On the other hand, by some elementary calculations, we have

$$\begin{aligned} \Big| \sum_{t=1}^T b_t' X_t' \boldsymbol{\Lambda}\boldsymbol{\Lambda}' \varepsilon_t \Big| &\le \sum_{t=1}^T \big| b_t' X_t' \boldsymbol{\Lambda}\boldsymbol{\Lambda}' \varepsilon_t \big| \le \max_{1\le t\le T} \big\| X_t' \boldsymbol{\Lambda} \big\| \sum_{t=1}^T \big\|b_t\big\| \big\|\boldsymbol{\Lambda}'\varepsilon_t\big\| \\ &\le \max_{1\le t\le T} \big\| X_t' \boldsymbol{\Lambda} \big\| \Big(\sum_{t=1}^T \big\|b_t\big\|^2 \Big)^{1/2} \Big(\sum_{t=1}^T \big\|\boldsymbol{\Lambda}'\varepsilon_t\big\|^2 \Big)^{1/2}. \end{aligned}$$

By the restriction on $\mathbf{\Lambda}$ and Assumption 1(ii), we have

$$\max_{1\le t\le T}\left\|X_t'\mathbf{\Lambda}\right\|^2=\max_{1\le t\le T}\mathrm{tr}\left(\mathbf{\Lambda}'X_tX_t'\mathbf{\Lambda}\right)\le\max_{1\le t\le T}\mu_{\max}\left(X_t'X_t\right)\|\mathbf{\Lambda}\|^2=O_P(N^2). \tag{C.2}$$

On the other hand, using $\frac{1}{N}\mathbf{\Lambda}'\mathbf{\Lambda}=\boldsymbol{I}_{R_0}$ and Assumption 1(iii), we have

$$\begin{aligned}\sum_{t=1}^T||\mathbf{\Lambda}'\varepsilon_t||^2 &= \sum_{t=1}^T\mathsf{Tr}(\mathbf{\Lambda}'\varepsilon_t\varepsilon_t'\mathbf{\Lambda})=\mathsf{Tr}(\mathbf{\Lambda}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{\Lambda})\\ &\le N\|\boldsymbol{\varepsilon}\|_{\mathrm{sp}}^2\,\mathsf{Tr}(\mathbf{\Lambda}'\mathbf{\Lambda}/N)=NR_0\|\boldsymbol{\varepsilon}\|_{\mathrm{sp}}^2=O_P\left(N(N+T)\right).\end{aligned} \tag{C.3}$$

It follows that

$$\Big|\sum_{t=1}^T b_t'X_t'\mathbf{\Lambda}\mathbf{\Lambda}'\varepsilon_t\Big|=O_P\big(p^{1/2}(N^2T^{1/2}+N^{3/2}T)\big), \tag{C.4}$$

as $\|\boldsymbol{b}\|\le C(pT)^{1/2}$. Then, by (C.1) and (C.4), we can complete the proof of (i).

(ii) By the definition of $\boldsymbol{M}_{\mathbf{\Lambda}}$ and noting that $\frac{1}{N}\mathbf{\Lambda}'\mathbf{\Lambda}=\boldsymbol{I}_{R_0}$, we have

$$\frac{1}{NT}\sum_{t=1}^T f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\boldsymbol{M}_{\mathbf{\Lambda}}\varepsilon_t=\frac{1}{NT}\sum_{t=1}^T f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\varepsilon_t-\frac{1}{N^2T}\sum_{t=1}^T f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}\mathbf{\Lambda}'\varepsilon_t.$$

By Assumptions 1(i) and (iii), we readily have

$$\Big|\sum_{t=1}^T f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\varepsilon_t\Big|=\Big(\sum_{t=1}^T\|f_t^{0\prime}\|^2\Big)^{1/2}\cdot\Big(\sum_{t=1}^T\|\mathbf{\Lambda}^{0\prime}\varepsilon_t\|^2\Big)^{1/2}=O_P(\sqrt{N}T). \tag{C.5}$$

On the other hand, as in the proof of (C.4) above we can show

$$\Big|\sum_{t=1}^T f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}\mathbf{\Lambda}'\varepsilon_t\Big|=O_P(N^2T^{1/2}+N^{3/2}T). \tag{C.6}$$

We then complete the proof of (ii) by using (C.5) and (C.6).

(iii) As $\frac{1}{N}\mathbf{\Lambda}'\mathbf{\Lambda}=\boldsymbol{I}_{R_0}$, we have $\frac{1}{NT}\sum_{t=1}^T\varepsilon_t'\boldsymbol{P}_{\mathbf{\Lambda}}\varepsilon_t=\frac{1}{N^2T}\sum_{t=1}^T\varepsilon_t'\mathbf{\Lambda}\mathbf{\Lambda}'\varepsilon_t$, which together with (C.3), completes the proof of (iii).

(iv) Using Assumption 1(iii) and the fact $\frac{1}{N}\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\xrightarrow{P}\mathbf{\Sigma}_{\mathbf{\Lambda}}$ under Assumption 1(i), we have

$$\begin{aligned}\Big|\frac{1}{NT}\sum_{t=1}^T\varepsilon_t'\boldsymbol{P}_{\mathbf{\Lambda}^0}\varepsilon_t\Big| &\le \frac{1}{N}\Big\|\Big(\frac{1}{N}\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\Big)^+\Big\|\cdot\frac{1}{NT}\sum_{t=1}^T\left\|\mathbf{\Lambda}^{0\prime}\varepsilon_t\right\|^2\\ &= O_P(N^{-1})\cdot O_P(1)\cdot O_P(1)=O_P(N^{-1}),\end{aligned} \tag{C.7}$$

which completes the proof of (iv).

We has thus completed the proof of Lemma C.1. ■

Lemma C.2 *Suppose that Assumption 1 in Appendix A holds and $pN^{-1/2}+p^{1/2}T^{-1/2}=o(1)$. Let $\dot{\boldsymbol{\beta}}=(\dot{\beta}_1',...,\dot{\beta}_T')'$ and $\dot{\boldsymbol{\Lambda}}=(\dot{\lambda}_1',...,\dot{\lambda}_N')'$ be the preliminary estimates of $\boldsymbol{\beta}^0$ and $\boldsymbol{\Lambda}^0$ which minimize $\hat{Q}_{NT}(\boldsymbol{\beta},\boldsymbol{\Lambda})$, the first term of the objective function defined in (2.4). Then*

$$\frac{1}{T}\sum_{t=1}^{T}\|\dot{\beta}_t-\beta_t^0\|^2=O_P\left(pN^{-1/2}+p^{1/2}T^{-1/2}\right)=o_P(1).$$

Proof of Lemma C.2. The proof of this lemma is similar to that of Theorem 3.1 in Appendix B of the main document. Notice that

$$\hat{Q}_{NT}(\boldsymbol{\beta},\boldsymbol{\Lambda})=\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{N}\big(Y_t-X_t\beta_t\big)'\boldsymbol{M}_{\boldsymbol{\Lambda}}\big(Y_t-X_t\beta_t\big)\right]\equiv\frac{1}{T}\sum_{t=1}^{T}\hat{Q}_{NT,t}(\beta_t,\boldsymbol{\Lambda}) \tag{C.8}$$

and

$$Y_t-X_t\dot{\beta}_t=X_t(\beta_t^0-\dot{\beta}_t)+\boldsymbol{\Lambda}^0f_t^0+\varepsilon_t. \tag{C.9}$$

Then, by (C.8) and (C.9) and using the fact that $\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\boldsymbol{\Lambda}^0=\mathbf{0}$, we have

$$\begin{aligned}
& Q_{NT}(\dot{\boldsymbol{\beta}},\dot{\boldsymbol{\Lambda}})-Q_{NT}(\boldsymbol{\beta}^0,\boldsymbol{\Lambda}^0)\\
= & \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\Big[\big(Y_t-X_t\dot{\beta}_t\big)'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\big(Y_t-X_t\dot{\beta}_t\big)-\big(Y_t-X_t\beta_t^0\big)'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\big(Y_t-X_t\beta_t^0\big)\Big]\\
= & \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\Big[\big(\dot{\beta}_t-\beta_t^0\big)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}X_t\big(\dot{\beta}_t-\beta_t^0\big)-2\big(\dot{\beta}_t-\beta_t^0\big)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0+f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0\Big]\\
& +\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\Big[-2\big(\dot{\beta}_t-\beta_t^0\big)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\varepsilon_t+2f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\varepsilon_t-\varepsilon_t'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}}\varepsilon_t+\varepsilon_t'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\varepsilon_t\Big].
\end{aligned} \tag{C.10}$$

By Lemma C.1 above, we can prove that

$$\frac{1}{NT}\sum_{t=1}^{T}\Big[-2\big(\dot{\beta}_t-\beta_t^0\big)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\varepsilon_t+2f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\varepsilon_t-\varepsilon_t'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}}\varepsilon_t+\varepsilon_t'\boldsymbol{P}_{\boldsymbol{\Lambda}^0}\varepsilon_t\Big]=O_P\big(pN^{-1/2}+p^{1/2}T^{-1/2}\big). \tag{C.11}$$

Let $\dot{\boldsymbol{d}}_\beta=\dot{\boldsymbol{\beta}}-\boldsymbol{\beta}^0$ and $\dot{\boldsymbol{d}}_\Lambda=\frac{1}{N^{1/2}}\mathsf{vec}(\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0)$ where $\mathsf{vec}(\cdot)$ denotes the vectorization of a matrix. Define

$$\begin{aligned}
\dot{\boldsymbol{A}} &= \frac{1}{N}\mathsf{diag}\big(X_1'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}X_1,...,X_T'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}X_T\big),\quad \dot{\boldsymbol{B}}=(\boldsymbol{F}^{0\prime}\boldsymbol{F}^0)\otimes\boldsymbol{I}_N,\ \text{and}\\
\dot{\boldsymbol{C}} &= \frac{1}{N^{1/2}}\big[f_1^0\otimes\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}X_1,...,f_T^0\otimes\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}X_T\big],
\end{aligned}$$

where $\otimes$ denotes the Kronecker product. It is easy to verify that

$$\frac{1}{NT}\sum_{t=1}^{T}\left(\dot{\beta}_t-\beta_t^0\right)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}X_t\left(\dot{\beta}_t-\beta_t^0\right)=\frac{1}{T}\dot{\boldsymbol{d}}_\beta'\dot{\boldsymbol{A}}\dot{\boldsymbol{d}}_\beta,$$
$$\frac{1}{NT}\sum_{t=1}^{T}\left(\dot{\beta}_t-\beta_t^0\right)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0=\frac{1}{NT}\sum_{t=1}^{T}\mathsf{Tr}\Big\{\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0\left(\dot{\beta}_t-\beta_t^0\right)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\Big\}=\frac{1}{T}\dot{\boldsymbol{d}}_\Lambda'\dot{\boldsymbol{C}}\dot{\boldsymbol{d}}_\beta,$$

and

$$\frac{1}{NT}\sum_{t=1}^{T}f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0=\frac{1}{NT}\sum_{t=1}^{T}\mathsf{Tr}\Big(\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\Big)=\frac{1}{T}\dot{\boldsymbol{d}}_\Lambda'\dot{\boldsymbol{B}}\dot{\boldsymbol{d}}_\Lambda,$$

where we have used the following fact on matrix calculation that $\mathsf{Tr}\left(\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\right)=\mathsf{vec}'\left(\boldsymbol{A}_1\right)\left(\boldsymbol{A}_2\otimes\boldsymbol{I}_k\right)\mathsf{vec}\left(\boldsymbol{A}_3\right)$ and that $\mathsf{Tr}\left(\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_3\boldsymbol{A}_4\right)=\mathsf{vec}'\left(\boldsymbol{A}_1\right)\left(\boldsymbol{A}_2\otimes\boldsymbol{A}_4'\right)\mathsf{vec}\left(\boldsymbol{A}_3'\right)$ with k being the size of the column vectors in $\boldsymbol{A}_3$ (in the first equation). With the above notations, we may show that

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\Big[\left(\dot{\beta}_t-\beta_t^0\right)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}X_t\left(\dot{\beta}_t-\beta_t^0\right)-2\left(\dot{\beta}_t-\beta_t^0\right)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0+f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}}\boldsymbol{\Lambda}^0f_t^0\Big]\\ =\;&\frac{1}{T}\left(\dot{\boldsymbol{d}}_\beta'\dot{\boldsymbol{A}}\dot{\boldsymbol{d}}_\beta-2\dot{\boldsymbol{d}}_\Lambda'\dot{\boldsymbol{C}}\dot{\boldsymbol{d}}_\beta+\dot{\boldsymbol{d}}_\Lambda'\dot{\boldsymbol{B}}\dot{\boldsymbol{d}}_\Lambda\right)=\frac{1}{T}\left(\dot{\boldsymbol{d}}_\beta'\dot{\boldsymbol{D}}\dot{\boldsymbol{d}}_\beta+\dot{\boldsymbol{d}}_*'\dot{\boldsymbol{B}}\dot{\boldsymbol{d}}_*\right),\end{aligned}$$

where $\dot{\boldsymbol{D}}=\dot{\boldsymbol{A}}-\dot{\boldsymbol{C}}'\dot{\boldsymbol{B}}^+\dot{\boldsymbol{C}}$ and $\dot{\boldsymbol{d}}_*=\dot{\boldsymbol{d}}_\Lambda-\dot{\boldsymbol{B}}^+\dot{\boldsymbol{C}}\dot{\boldsymbol{d}}_\beta$. By Assumption 1(i), we may show that the minimum eigenvalue of $\frac{1}{T}\dot{\boldsymbol{B}}$ is bounded away from zero w.p.a.1, i.e., there exists a positive constant c_4 such that $\mu_{\min}\left(\dot{\boldsymbol{B}}/T\right)>c_4$ w.p.a.1. Using a decomposition similar to (B.8) in Appendix B, we can readily show that $\mu_{\max}\left(\dot{\boldsymbol{C}}'\dot{\boldsymbol{C}}/T\right)=o_P(1)$. By Assumption 1(ii), we can also show that the minimum eigenvalue of $\dot{\boldsymbol{A}}$ is bounded away from zero w.p.a.1, i.e., there exists a positive constant c_x (defined in Assumption 1(ii)) such that $\mu_{\min}\left(\dot{\boldsymbol{A}}\right)>c_x$ w.p.a.1. Hence, we have proved that the matrix $\dot{\boldsymbol{D}}$ is asymptotically positive definite as its minimum eigenvalue is positive and bounded away from zero w.p.a.1.

Note that

$$\frac{1}{T}\left(\dot{\boldsymbol{d}}_\beta'\dot{\boldsymbol{D}}\dot{\boldsymbol{d}}_\beta+\dot{\boldsymbol{d}}_*'\dot{\boldsymbol{B}}\dot{\boldsymbol{d}}_*\right)+O_P\left(pN^{-1/2}+p^{1/2}T^{-1/2}\right)\leq Q_{NT}\left(\dot{\boldsymbol{\beta}},\dot{\boldsymbol{\Lambda}}\right)-Q_{NT}\left(\boldsymbol{\beta}^0,\boldsymbol{\Lambda}^0\right)\leq 0,\quad\text{(C.12)}$$

$\dot{\boldsymbol{d}}_*'\dot{\boldsymbol{B}}\dot{\boldsymbol{d}}_*$ is asymptotically nonnegative, and $\dot{\boldsymbol{d}}_\beta'\dot{\boldsymbol{D}}\dot{\boldsymbol{d}}_\beta\geq c_5\|\dot{\boldsymbol{d}}_\beta\|^2$ where c_5 is a positive constant. It follows that $\frac{1}{T}\|\dot{\boldsymbol{d}}_\beta\|^2=\frac{1}{T}\sum_{t=1}^{T}\|\dot{\beta}_t-\beta_t^0\|^2=O_P(pN^{-1/2}+p^{1/2}T^{-1/2})=o_P(1)$, completing the proof of Lemma C.2. ∎

Lemma C.3 *Suppose that Assumption 1 in Appendix A holds and $pN^{-1/2}+p^{1/2}T^{-1/2}=o(1)$. Let $\dot{\boldsymbol{H}}\equiv\dot{\boldsymbol{H}}_{NT}=\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\dot{\boldsymbol{\Lambda}}\right)\dot{\boldsymbol{V}}_{NT}^+$, where $\dot{\boldsymbol{V}}_{NT}$ is analogously defined as $\boldsymbol{V}_{NT}$ in (2.7) with $\hat{\beta}_t$ replaced by $\dot{\beta}_t$. Denote $\dot{\eta}_{NT}=\frac{1}{T}\sum_{t=1}^{T}\|\dot{\beta}_t-\beta_t^0\|^2$. Then we have*

(i) $\frac{1}{N}\big\|\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\big\|^2=O_P\big(\delta_{NT}^{-2}+\dot{\eta}_{NT}\big)$,
(ii) $\frac{1}{N}\big(\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\big)'\mathbf{\Lambda}^0\dot{\boldsymbol{H}}=O_P\big(\delta_{NT}^{-2}+\dot{\eta}_{NT}^{1/2}\big)$,
(iii) $\frac{1}{N}\big(\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\big)'\dot{\mathbf{\Lambda}}=O_P\big(\delta_{NT}^{-2}+\dot{\eta}_{NT}^{1/2}\big)$,
(iv) $\frac{1}{N}\big(\dot{\mathbf{\Lambda}}'\dot{\mathbf{\Lambda}}-\dot{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\big)=O_P\big(\delta_{NT}^{-2}+\dot{\eta}_{NT}^{1/2}\big)$,
(v) $\big\|\boldsymbol{P}_{\dot{\mathbf{\Lambda}}}-\boldsymbol{P}_{\mathbf{\Lambda}^0\dot{\boldsymbol{H}}}\big\|=O_P\big(\delta_{NT}^{-2}+\dot{\eta}_{NT}^{1/2}\big)$,
(vi) $\frac{1}{NT}\sum_{s=1}^{T}(\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}})'\varepsilon_s\gamma_s'=O_P\big(\delta_{NT}^{-2}+\dot{\eta}_{NT}^{1/2}\big)$ *with* $\gamma_s=1$ *or* f_s^0*, and*
(vii) $\frac{1}{NT}\sum_{s=1}^{T}\|(\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}})'\varepsilon_s\|^2=O_P\big((1+NT^{-1})(\delta_{NT}^{-2}+\dot{\eta}_{NT})\big)$.

Proof of Lemma C.3. **(i)** By (2.7) and (C.9) and letting $d_t=\dot{\beta}_t-\beta_t^0$, we have

$$\begin{aligned}
&\dot{\mathbf{\Lambda}}\dot{\boldsymbol{V}}_{NT}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\dot{\boldsymbol{V}}_{NT}\\
=&\Big[\frac{1}{NT}\sum_{t=1}^{T}\big(Y_t-X_t\dot{\beta}_t\big)(Y_t-X_t\dot{\beta}_t)\Big]\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\dot{\boldsymbol{V}}_{NT}\\
=&\Big\{\frac{1}{NT}\sum_{t=1}^{T}\big[-X_td_t+\mathbf{\Lambda}^0f_t^0+\varepsilon_t\big]\big[-X_td_t+\mathbf{\Lambda}^0f_t^0+\varepsilon_t\big]'\Big\}\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\dot{\boldsymbol{V}}_{NT}\\
=&\frac{1}{NT}\sum_{t=1}^{T}X_td_td_t'X_t'\dot{\mathbf{\Lambda}}-\frac{1}{NT}\sum_{t=1}^{T}X_td_tf_t^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}-\frac{1}{NT}\sum_{t=1}^{T}X_td_t\varepsilon_t'\dot{\mathbf{\Lambda}}-\frac{1}{NT}\sum_{t=1}^{T}\mathbf{\Lambda}^0f_t^0d_t'X_t'\dot{\mathbf{\Lambda}}\\
&+\frac{1}{NT}\sum_{t=1}^{T}\mathbf{\Lambda}^0f_t^0\varepsilon_t'\dot{\mathbf{\Lambda}}-\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_td_t'X_t'\dot{\mathbf{\Lambda}}+\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_tf_t^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}+\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t\varepsilon_t'\dot{\mathbf{\Lambda}}\\
\equiv&\sum_{j=1}^{8}\dot{u}_{NT,j}.
\end{aligned}\tag{C.13}$$

Noting that $\mathsf{Tr}\,(AB)\leq\mathsf{Tr}\,(A)\,\mathsf{Tr}\,(B)$ for conformable positive semidefinite matrices A and B, $\|\dot{\mathbf{\Lambda}}\|=O_P(N^{1/2})$ and $\max_{1\leq t\leq T}\mu_{\max}^2\,(X_t'X_t/N)=O_P\,(1)$ by Assumption 1(ii), we have

$$\begin{aligned}
\|\dot{u}_{NT,1}\|^2&=\frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}(X_td_td_t'X_t'\dot{\mathbf{\Lambda}}\dot{\mathbf{\Lambda}}'X_sd_sd_s'X_s')\\
&\leq\big\|\dot{\mathbf{\Lambda}}\big\|^2\left\{\frac{1}{NT}\sum_{t=1}^{T}\mathsf{Tr}(X_td_td_t'X_t')\right\}^2=\big\|\dot{\mathbf{\Lambda}}\big\|^2\left\{\frac{1}{NT}\sum_{t=1}^{T}d_t'X_t'X_td_t\right\}^2\\
&\leq\big\|\dot{\mathbf{\Lambda}}\big\|^2\left[\max_{1\leq t\leq T}\mu_{\max}^2\left(X_t'X_t/N\right)\right]\left\{\frac{1}{T}\sum_{t=1}^{T}\|d_t\|^2\right\}^2=O_P(N\dot{\eta}_{NT}^2).
\end{aligned}\tag{C.14}$$

Noting that $\mathsf{Tr}\,(AB) \leq \mathsf{Tr}\,(AA')^{1/2}\,\mathsf{Tr}\,(BB')^{1/2}$ for conformable matrices A and B, we have

$$\begin{aligned}
\|\dot{u}_{NT,2}\|^2 &= \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}(X_t d_t f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}\dot{\mathbf{\Lambda}}'\mathbf{\Lambda}^0 f_s^0 d_s' X_s') \\
&\leq \left\|\dot{\mathbf{\Lambda}}\right\|^2 \mu_{\max}(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0)\frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}(X_t d_t f_t^{0\prime} f_s^0 d_s' X_s') \\
&\leq \frac{1}{N}\left\|\dot{\mathbf{\Lambda}}\right\|^2 \mu_{\max}(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0/N)\left(\frac{1}{T}\sum_{t=1}^{T}\left\{\mathsf{Tr}(f_t^0 d_t' X_t' X_t d_t f_t^{0\prime})\right\}^{1/2}\right)^2 \\
&\leq \left\|\dot{\mathbf{\Lambda}}\right\|^2 \mu_{\max}(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0/N)\left[\max_{1\leq t\leq T}\mu_{\max}\left(X_t'X_t/N\right)\right]\left(\frac{1}{T}\sum_{t=1}^{T}\|d_t\|\left\|f_t^0\right\|\right)^2 \\
&= O_P\left(N\right)O_P\left(1\right)O_P\left(1\right)\frac{1}{T}\sum_{t=1}^{T}\|d_t\|^2\frac{1}{T}\sum_{t=1}^{T}\left\|f_t^0\right\|^2 = O_P\left(N\dot{\eta}_{NT}\right), \qquad \text{(C.15)}
\end{aligned}$$

and analogously

$$\|\dot{u}_{NT,4}\|^2 = O_P\left(N\left(\frac{1}{T}\sum_{t=1}^{T}\|\dot{\beta}_t-\beta_t^0\|^2\right)\right) = O_P\left(N\dot{\eta}_{NT}\right). \qquad \text{(C.16)}$$

Noting that $\sum_{t=1}^{T}\|\varepsilon_t\|^2 = O_P(NT)$ by Assumption 1(iii) and $\max_{1\leq t\leq T}\mu_{\max}\left(X_t'X_t/N\right) = O_P\big(1\big)$ by Assumption 1(ii), we can show that

$$\begin{aligned}
\|\dot{u}_{NT,3}\|^2 &= \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}(X_t d_t \varepsilon_t'\dot{\mathbf{\Lambda}}\dot{\mathbf{\Lambda}}'\varepsilon_s d_s' X_s') \leq \left\|\dot{\mathbf{\Lambda}}\right\|^2\frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}(X_t d_t \varepsilon_t'\varepsilon_s d_s' X_s') \\
&\leq \left\|\dot{\mathbf{\Lambda}}\right\|^2\left\{\frac{1}{NT}\sum_{t=1}^{T}\left\{\mathsf{Tr}(\varepsilon_t d_t' X_t' X_t d_t \varepsilon_t')\right\}^{1/2}\right\}^2 \\
&\leq \frac{1}{N}\left\|\dot{\mathbf{\Lambda}}\right\|^2\left[\max_{1\leq t\leq T}\mu_{\max}\left(X_t'X_t/N\right)\right]\left\{\frac{1}{T}\sum_{t=1}^{T}\|\varepsilon_t\|\,\|d_t\|\right\}^2 \\
&\leq O_P\left(1\right)\frac{1}{T}\sum_{t=1}^{T}\|\varepsilon_t\|^2\,\frac{1}{T}\sum_{t=1}^{T}\|d_t\|^2 = O_P\left(N\dot{\eta}_{NT}\right) \qquad \text{(C.17)}
\end{aligned}$$

and analogously

$$\|\dot{u}_{NT,6}\|^2 = O_P\left(\frac{N}{T}\sum_{t=1}^{T}\|\dot{\beta}_t-\beta_t^0\|^2\right) = O_P\left(N\dot{\eta}_{NT}\right). \qquad \text{(C.18)}$$

The analysis of the remaining three terms is similar to the proof of Theorem 1 in Bai and Ng (2002) by switching the roles of f_t and λ_i. For $\dot{u}_{NT,5}$, using the fact that $\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0 = O_P(N)$,

$\|\dot{\mathbf{\Lambda}}\| = O_P(N^{1/2})$ and Assumptions 1(iii) and (iv), we can prove that

$$\begin{aligned}
\|\dot{u}_{NT,5}\|^2 &= \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}\Big(\mathbf{\Lambda}^0 f_t^0 \varepsilon_t' \dot{\mathbf{\Lambda}}\dot{\mathbf{\Lambda}}' \varepsilon_s f_s^{0\prime}\mathbf{\Lambda}^{0\prime}\Big) = \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}\Big(f_t^0 \varepsilon_t' \dot{\mathbf{\Lambda}}\dot{\mathbf{\Lambda}}' \varepsilon_s f_s^{0\prime}\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\Big) \\
&= O_P\left(\frac{1}{NT^2}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}\sum_{k=1}^{N}\varepsilon_{it}\varepsilon_{ks}\dot{\lambda}_i'\dot{\lambda}_k f_t^0 f_s^{0\prime}\right\|\right) \\
&= O_P\left(\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{k=1}^{N}\left|\dot{\lambda}_i'\dot{\lambda}_k\right|\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}\varepsilon_{ks}f_t^0 f_s^{0\prime}\right\|\right) \\
&= O_P\left(\frac{1}{NT^2}\left(\sum_{i=1}^{N}\sum_{k=1}^{N}\|\dot{\lambda}_i\|^2\|\dot{\lambda}_k\|^2\right)^{1/2}\left(\sum_{i=1}^{N}\sum_{k=1}^{N}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}\varepsilon_{ks}f_t^0 f_s^{0\prime}\right\|^2\right)^{1/2}\right) \\
&= O_P\left(\frac{1}{T^2}\left(\sum_{i=1}^{N}\sum_{k=1}^{N}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}\varepsilon_{ks}f_t^0 f_s^{0\prime}\right\|^2\right)^{1/2}\right) = O_P(N/T), \qquad \text{(C.19)}
\end{aligned}$$

and

$$\begin{aligned}
\|\dot{u}_{NT,7}\|^2 &= \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}\left(\varepsilon_t f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}\dot{\mathbf{\Lambda}}'\mathbf{\Lambda}^0 f_s^0 \varepsilon_s'\right) = \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathsf{Tr}\left(\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}\dot{\mathbf{\Lambda}}'\mathbf{\Lambda}^0 f_s^0 \varepsilon_s'\varepsilon_t f_t^{0\prime}\right) \\
&= O_P\left(\frac{1}{T^2}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T} f_s^0\varepsilon_s'\varepsilon_t f_t^{0\prime}\right\|\right) = O_P(N/T). \qquad \text{(C.20)}
\end{aligned}$$

By the assumption that $\max_{1\le i,j\le N}\mathsf{E}\left[\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{it}\varepsilon_{js}\varepsilon_t'\varepsilon_s\right\|^2\right] = O(N^2T^2+T^2)$ in Assumption 1(iii), we can similarly prove

$$\|\dot{u}_{NT,8}\|^2 = O_P(N/T). \qquad \text{(C.21)}$$

By (C.13)–(C.21), we can prove that

$$\frac{1}{N}\left\|\dot{\mathbf{\Lambda}}\dot{\boldsymbol{V}}_{NT} - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}\dot{\boldsymbol{V}}_{NT}\right\|^2 = O_P(\delta_{NT}^{-2} + \dot{\eta}_{NT}). \qquad \text{(C.22)}$$

Premultiplying (C.13) by $\dot{\mathbf{\Lambda}}'$, and using the identification restriction on $\dot{\mathbf{\Lambda}}$: $\frac{1}{N}\dot{\mathbf{\Lambda}}'\dot{\mathbf{\Lambda}} = \boldsymbol{I}_{R_0}$, (C.22) and Lemma C.2, we may show that

$$\dot{\boldsymbol{V}}_{NT} - \left(\frac{1}{N}\dot{\mathbf{\Lambda}}'\mathbf{\Lambda}^0\right)\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{N}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}\right) = o_P(1). \qquad \text{(C.23)}$$

Furthermore, applying (C.12) in the proof of Lemma C.2 and noting that the matrix $\dot{\boldsymbol{B}}$ is positive definite, we can show that

$$\frac{1}{N}\mathbf{\Lambda}^{0\prime}\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}\mathbf{\Lambda}^0 = \frac{1}{N}\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0 - \left(\frac{1}{N}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}\right)\left(\frac{1}{N}\dot{\mathbf{\Lambda}}'\mathbf{\Lambda}^0\right) = o_P(1),$$

which together with Assumption 1(i), implies that $\frac{1}{N}\dot{\mathbf{\Lambda}}'\mathbf{\Lambda}^0$ is asymptotically invertible and thus $\dot{\boldsymbol{V}}_{NT}$ is also asymptotically invertible. We can then complete the proof of (i) by using this fact and (C.22).

(ii) Observe that by (C.13)

$$\frac{1}{N}(\dot{\mathbf{\Lambda}} - \mathbf{\Lambda}^0\dot{\boldsymbol{H}})'\mathbf{\Lambda}^0\dot{\boldsymbol{H}} = \frac{1}{N}\sum_{j=1}^{8}\dot{\boldsymbol{V}}_{NT}^{+}\dot{u}_{NT,j}'\mathbf{\Lambda}^0\dot{\boldsymbol{H}} \equiv \frac{1}{N}\sum_{j=1}^{8}\dot{u}_{NT,j}^{*}. \tag{C.24}$$

By Assumption 1(i) and (C.14), we can readily prove

$$\frac{1}{N}\|\dot{u}_{NT,1}^{*}\| \leq \left(\frac{1}{N^{1/2}}\|\dot{u}_{NT,1}\|\right)\cdot\|\dot{\boldsymbol{V}}_{NT}^{+}\|\cdot\left(\frac{1}{N^{1/2}}\|\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\|\right) = O_P\left(\dot{\eta}_{NT}\right). \tag{C.25}$$

Analogously, by (C.15) and (C.16), we can prove that

$$\frac{1}{N}\|\dot{u}_{NT,2}^{*}\| = O_P\left(\dot{\eta}_{NT}^{1/2}\right) \quad \text{and} \quad \frac{1}{N}\|\dot{u}_{NT,4}^{*}\| = O_P\left(\dot{\eta}_{NT}^{1/2}\right). \tag{C.26}$$

For $\dot{u}_{NT,3}^{*}$, by the definition of $\dot{u}_{NT,3}$, we have

$$\begin{aligned}
-\dot{u}_{NT,3}^{*} &= -\dot{\boldsymbol{V}}_{NT}^{+}\dot{u}_{NT,3}'\mathbf{\Lambda}^0\dot{\boldsymbol{H}} = \frac{1}{NT}\dot{\boldsymbol{V}}_{NT}^{+}\sum_{t=1}^{T}\dot{\mathbf{\Lambda}}'\varepsilon_t d_t' X_t'\mathbf{\Lambda}^0\dot{\boldsymbol{H}} \\
&= \frac{1}{NT}\dot{\boldsymbol{V}}_{NT}^{+}\sum_{t=1}^{T}\dot{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime}\varepsilon_t d_t' X_t'\mathbf{\Lambda}^0\dot{\boldsymbol{H}} + \frac{1}{NT}\dot{\boldsymbol{V}}_{NT}^{+}\sum_{t=1}^{T}(\dot{\mathbf{\Lambda}} - \mathbf{\Lambda}^0\dot{\boldsymbol{H}})'\varepsilon_t d_t' X_t'\mathbf{\Lambda}^0\dot{\boldsymbol{H}} \\
&\equiv \dot{u}_{NT,3a}^{*} + \dot{u}_{NT,3b}^{*}.
\end{aligned} \tag{C.27}$$

By the Cauchy-Schwarz inequality and Assumptions 1(ii) and (iii), we have

$$\|\dot{u}_{NT,3a}^{*}\| \leq \frac{C}{T}\sum_{t=1}^{T}\|\mathbf{\Lambda}^{0\prime}\varepsilon_t d_t'\| \leq C\left(\frac{1}{T}\sum_{t=1}^{T}\|\mathbf{\Lambda}^{0\prime}\varepsilon_t\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\|d_t\|^2\right)^{1/2} = O_P\left((N\dot{\eta}_{NT})^{1/2}\right). \tag{C.28}$$

Similarly, with the help of Lemma C.3(i), we can also prove that

$$\|\dot{u}_{NT,3b}^{*}\| = O_P\left(N\dot{\eta}_{NT} + N\delta_{NT}^{-1}\dot{\eta}_{NT}^{1/2}\right). \tag{C.29}$$

By (C.27)–(C.29), we have

$$\frac{1}{N}\|\dot{u}_{NT,3}^{*}\| = O_P\left(\dot{\eta}_{NT} + \delta_{NT}^{-1}\dot{\eta}_{NT}^{1/2}\right). \tag{C.30}$$

Similarly, we can also show that

$$\frac{1}{N}\|\dot{u}_{NT,6}^{*}\| = O_P\left(\dot{\eta}_{NT} + \delta_{NT}^{-1}\dot{\eta}_{NT}^{1/2}\right). \tag{C.31}$$

For $\dot{u}^*_{NT,5}$, by the definition of $\dot{u}_{NT,5}$, we have

$$\begin{aligned}\dot{u}^*_{NT,5} &= \frac{1}{NT}\dot{\boldsymbol{V}}^+_{NT}\sum_{t=1}^T \dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\varepsilon_t f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}} + \frac{1}{NT}\dot{\boldsymbol{V}}^+_{NT}\sum_{t=1}^T\left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)'\varepsilon_t f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}} \\ &\equiv \dot{u}^*_{NT,5a} + \dot{u}^*_{NT,5b}.\end{aligned} \tag{C.32}$$

By Assumptions 1(i) and (iii), we have

$$\|\dot{u}^*_{NT,5a}\| \le C\frac{1}{T}\Big\|\sum_{t=1}^T \boldsymbol{\Lambda}^{0\prime}\varepsilon_t f_t^{0\prime}\Big\| = O_P\Big(\frac{1}{T}\big\|\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0\big\|\Big) = O_P\left(N^{1/2}T^{-1/2}\right). \tag{C.33}$$

Using Lemma C.3(i), we can also prove that

$$\|\dot{u}^*_{NT,5b}\| = O_P\left(N\dot{\eta}_{NT} + N\delta_{NT}^{-2}\right). \tag{C.34}$$

By (C.32)–(C.34), we have

$$\frac{1}{N}\|\dot{u}^*_{NT,5}\| = O_P\left(\dot{\eta}_{NT} + \delta_{NT}^{-2}\right). \tag{C.35}$$

Noting that $\dot{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0 = O_P(N)$ and using the assumption $\mathsf{E}\left[\big\|\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0\big\|^2\right] = O(NT)$ in Assumption 1(iii), we can also show that

$$\frac{1}{N}\|\dot{u}^*_{NT,7}\| = O_P\left(\dot{\eta}_{NT} + \delta_{NT}^{-2}\right) \quad \text{and} \quad \frac{1}{N}\|\dot{u}^*_{NT,8}\| = O_P\left(\dot{\eta}_{NT} + \delta_{NT}^{-2}\right). \tag{C.36}$$

By (C.24)–(C.26), (C.30), (C.31), (C.35) and (C.36), we can complete the proof of (ii).

(iii) and **(iv)** The proofs of (iii) and (iv) can be completed by using the results in Lemmas C.3(i) and (ii).

(v) Note that

$$\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}} = \dot{\boldsymbol{\Lambda}}(\dot{\boldsymbol{\Lambda}}'\dot{\boldsymbol{\Lambda}})^+\dot{\boldsymbol{\Lambda}}' - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime} \equiv \sum_{j=1}^7 \dot{v}_{NT,j}, \tag{C.37}$$

where

$$\begin{aligned}
\dot{v}_{NT,1} &= \left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)\left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)', \\
\dot{v}_{NT,2} &= \left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)\left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}, \\
\dot{v}_{NT,3} &= \left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)\left[\left(\dot{\boldsymbol{\Lambda}}'\dot{\boldsymbol{\Lambda}}\right)^+ - \left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\right]\left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)', \\
\dot{v}_{NT,4} &= \left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)\left[\left(\dot{\boldsymbol{\Lambda}}'\dot{\boldsymbol{\Lambda}}\right)^+ - \left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\right]\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}, \\
\dot{v}_{NT,5} &= \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)', \\
\dot{v}_{NT,6} &= \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\left[\left(\dot{\boldsymbol{\Lambda}}'\dot{\boldsymbol{\Lambda}}\right)^+ - \left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\right]\left(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)', \\
\dot{v}_{NT,7} &= \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\left[\left(\dot{\boldsymbol{\Lambda}}'\dot{\boldsymbol{\Lambda}}\right)^+ - \left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\right]\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}.
\end{aligned}$$

Using the results in Lemmas C.3(i) and (iv), we can prove (v).

(**vi**) The proof is analogous to that of part (ii) and thus omitted.

(**vii**) By Assumption 1(iii) and part (i),

$$
\begin{aligned}
\frac{1}{NT}\sum_{s=1}^{T}||(\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}})'\varepsilon_s||^2 &= \frac{1}{NT}\mathsf{Tr}\Big((\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}})'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'(\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}})\Big) \\
&\leq \frac{1}{T}\|\boldsymbol{\varepsilon}\|_{\mathrm{sp}}^2\cdot\frac{1}{N}\mathsf{Tr}\big((\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}})'(\dot{\mathbf{\Lambda}}-\mathbf{\Lambda}^0\dot{\boldsymbol{H}})\big) \\
&= O_P((1+NT^{-1})(\delta_{NT}^{-2}+\dot{\eta}_{NT})).
\end{aligned}
$$

We have thus completed the proof of Lemma C.3. ■

With the above three lemmas, we are ready to give the proof of Lemma B.1.

Proof of Lemma B.1. Let $\hat{Q}_{NT,t}(\beta_t,\mathbf{\Lambda})$ be defined as in (C.8), $\dot{\boldsymbol{\beta}}$ and $\dot{\mathbf{\Lambda}}$ be defined in Lemma C.2, and $\dot{\boldsymbol{H}}$ be defined in Lemma C.3. Note that

$$
Y_t - X_t\dot{\beta}_t = X_t(\beta_t^0-\dot{\beta}_t)+\dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+f_t^0+\big(\mathbf{\Lambda}^0-\dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+\big)f_t^0+\varepsilon_t. \tag{C.38}
$$

The preliminary estimate $\dot{\beta}_t$ which minimizes $\hat{Q}_{NT,t}(\beta_t,\mathbf{\Lambda})$ (with respect to β_t) satisfies that

$$
\Big(\frac{1}{N}X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}X_t\Big)(\dot{\beta}_t-\beta_t^0) = \frac{1}{N}X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}\varepsilon_t+\frac{1}{N}X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}\big(\mathbf{\Lambda}^0-\dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+\big)f_t^0, \tag{C.39}
$$

as $\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}\dot{\mathbf{\Lambda}}=\mathbf{0}$, where $\mathbf{0}$ is a null matrix or vector whose size may change from line to line.

We first consider the term $\frac{1}{N}X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}\varepsilon_t$. Notice that

$$
\frac{1}{N}X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}\varepsilon_t = \frac{1}{N}X_t'\boldsymbol{M}_{\mathbf{\Lambda}^0}\varepsilon_t+\frac{1}{N}X_t'(\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}-\boldsymbol{M}_{\mathbf{\Lambda}^0})\varepsilon_t. \tag{C.40}
$$

By the definition of $\boldsymbol{M}_{\mathbf{\Lambda}^0}$, we have

$$
\frac{1}{N}X_t'\boldsymbol{M}_{\mathbf{\Lambda}^0}\varepsilon_t = \frac{1}{N}X_t'\varepsilon_t-\frac{1}{N}X_t'\mathbf{\Lambda}^0(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0)^+\mathbf{\Lambda}^{0\prime}\varepsilon_t. \tag{C.41}
$$

By Assumption 1(iii), we can show that for each $1\leq t\leq T$

$$
\frac{1}{N}\|X_t'\varepsilon_t\| = O_P\left(p^{1/2}N^{-1/2}\right). \tag{C.42}
$$

By Assumptions 1(i)–(iii), we can show that for each $1\leq t\leq T$

$$
\|X_t'\mathbf{\Lambda}^0\|=O_P(N),\quad \|\mathbf{\Lambda}^{0\prime}\varepsilon_t\|=O_P(N^{1/2})\ \text{ and }\ \left(\frac{1}{N}\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\right)^+\xrightarrow{P}\mathbf{\Sigma}_\Lambda^+,
$$

which imply that

$$
\frac{1}{N}\|X_t'\mathbf{\Lambda}^0(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0)^+\mathbf{\Lambda}^{0\prime}\varepsilon_t\| = O_P\left(N^{-1/2}\right). \tag{C.43}
$$

Thus, by (C.41)–(C.43), we have

$$\frac{1}{N}\|X_t'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\varepsilon_t\| = O_P\left(p^{1/2}N^{-1/2}\right). \tag{C.44}$$

To derive the order of $X_t'(\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}} - \boldsymbol{M}_{\boldsymbol{\Lambda}^0})\varepsilon_t$, we need to investigate the term $\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}} - \boldsymbol{M}_{\boldsymbol{\Lambda}^0}$. By (C.37), we have

$$-(\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}} - \boldsymbol{M}_{\boldsymbol{\Lambda}^0}) = \dot{\boldsymbol{\Lambda}}\big(\dot{\boldsymbol{\Lambda}}'\dot{\boldsymbol{\Lambda}}\big)^+\dot{\boldsymbol{\Lambda}}' - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big)^+\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime} = \sum_{j=1}^{7}\dot{v}_{NT,j}. \tag{C.45}$$

We next show that

$$\frac{1}{N}\big\|X_t'\big(\sum_{j=1}^{7}\dot{v}_{NT,j}\big)\varepsilon_t\big\| = O_P\big(\delta_{NT}^{-1}\big). \tag{C.46}$$

To save the space, we only consider the case of $j=5$. Other cases can be studied similarly. For $X_t'\dot{v}_{NT,5}\varepsilon_t$, note that

$$\begin{aligned}
\dot{v}_{NT,5} &= \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big)^+\big(\dot{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big)', \\
&= \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big)^+\dot{\boldsymbol{V}}_{NT}^+\big(\dot{\boldsymbol{\Lambda}}\dot{\boldsymbol{V}}_{NT} - \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\dot{\boldsymbol{V}}_{NT}\big)', \\
&= \boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big)^+\dot{\boldsymbol{V}}_{NT}^+\big(\sum_{j=1}^{8}\dot{u}_{NT,j}\big)',
\end{aligned} \tag{C.47}$$

where $\dot{u}_{NT,j}$, $j=1,...,8$, are defined in the proof of Lemma C.3(i) above. By the fact that both $\dot{\boldsymbol{H}}$ and $\dot{\boldsymbol{V}}_{NT}$ are asymptotically invertible and similar to the proof of Lemma C.3(i), we readily prove that

$$\frac{1}{N}\left\|X_t'\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\left(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\dot{\boldsymbol{V}}_{NT}^+\left(\sum_{j=1}^{5}\dot{u}_{NT,j} + \dot{u}_{NT,8}\right)'\varepsilon_t\right\| = O_P\left(\delta_{NT}^{-2} + \delta_{NT}^{-1}\dot{\eta}_{NT}^{1/2}\right). \tag{C.48}$$

Meanwhile, by Assumptions 1(i)(ii) and noting that

$$\max_{1\le t\le T}\mathsf{E}\big[\sum_{s=1}^{T}\big|\varepsilon_s'\varepsilon_t\big|^2\big] = \max_{1\le t\le T}\mathsf{E}\big[\sum_{s=1}^{T}(\xi_{st}^*)^2\big] = O(N^2+NT)$$

by Assumption 1(iv), we can prove that

$$\begin{aligned}
&\frac{1}{N}\left\|X_t'\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big)^+\dot{\boldsymbol{V}}_{NT}^+\dot{u}_{NT,6}'\varepsilon_t\right\| \\
=\;& \frac{1}{N}\left\|X_t'\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big(\dot{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}\big)^+\dot{\boldsymbol{V}}_{NT}^+\left(\frac{1}{NT}\sum_{s=1}^{T}\varepsilon_s d_s' X_s'\dot{\boldsymbol{\Lambda}}\right)'\varepsilon_t\right\| \\
=\;& O_P\left(\frac{1}{N^2T}\left\|\sum_{s=1}^{T}\dot{\boldsymbol{\Lambda}}'X_s d_s\varepsilon_s'\varepsilon_t\right\|\right)
\end{aligned}$$

and

$$\begin{aligned}\frac{1}{N^2T}\left\|\sum_{s=1}^{T}\dot{\mathbf{\Lambda}}'X_s d_s\varepsilon_s'\varepsilon_t\right\| &\leq N^{-1/2}\left(\frac{1}{N^2T}\sum_{s=1}^{T}\left\|\dot{\mathbf{\Lambda}}'X_s d_s\right\|^2\right)^{1/2}\cdot\left(\frac{1}{NT}\sum_{s=1}^{T}\left\|\varepsilon_s'\varepsilon_t\right\|^2\right)^{1/2} \\ &= O_P\left(\delta_{NT}^{-1}\left(\frac{1}{T}\sum_{s=1}^{T}\|d_s\|^2\right)^{1/2}\right),\end{aligned}$$

which together with Lemma C.2, indicate that

$$\frac{1}{N}\left\|X_t'\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\left(\dot{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\dot{\mathbf{V}}_{NT}^+\dot{u}_{NT,6}'\varepsilon_t\right\| = O_P\left(\delta_{NT}^{-1}\dot{\eta}_{NT}^{1/2}\right). \tag{C.49}$$

Similarly, we can also show that

$$\begin{aligned}&\frac{1}{N}\left\|X_t'\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\left(\dot{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\dot{\mathbf{V}}_{NT}^+\dot{u}_{NT,7}'\varepsilon_t\right\| \\ =& \frac{1}{N}\left\|X_t'\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\left(\dot{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\right)^+\dot{\mathbf{V}}_{NT}^+\left(\frac{1}{NT}\sum_{s=1}^{T}\varepsilon_s f_s^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}\right)'\varepsilon_t\right\| = O_P(1)\,\frac{1}{NT}\left\|\sum_{s=1}^{T}f_s^0\varepsilon_s'\varepsilon_t\right\| \\ =& O_P(N^{-1/2})\left(\frac{1}{T}\sum_{s=1}^{T}\left\|f_s^0\right\|^2\right)^{1/2}\cdot\left(\frac{1}{NT}\sum_{s=1}^{T}\left\|\varepsilon_s'\varepsilon_t\right\|^2\right)^{1/2} = O_P\left(\delta_{NT}^{-1}\right).\end{aligned} \tag{C.50}$$

Then, by (C.48)–(C.50) and using the fact that $\dot{\eta}_{NT} = o_P(1)$ in Lemma C.2, we can readily prove that

$$\frac{1}{N}\left\|X_t'\dot{v}_{NT,5}\varepsilon_t\right\| = O_P\left(\delta_{NT}^{-1}\right). \tag{C.51}$$

Then we complete the proof of (C.46), which implies that

$$\frac{1}{N}\left\|X_t'(\boldsymbol{M}_{\dot{\mathbf{\Lambda}}} - \boldsymbol{M}_{\mathbf{\Lambda}^0})\varepsilon_t\right\| = O_P\left(\delta_{NT}^{-1}\right). \tag{C.52}$$

We next consider the term $\frac{1}{N}X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}(\mathbf{\Lambda}^0 - \dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+)f_t^0$. Note that

$$\frac{1}{N}X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}(\mathbf{\Lambda}^0-\dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+)f_t^0 = \frac{1}{N}X_t'\boldsymbol{M}_{\mathbf{\Lambda}^0\dot{\boldsymbol{H}}}(\mathbf{\Lambda}^0-\dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+)f_t^0+\frac{1}{N}X_t'(\boldsymbol{M}_{\dot{\mathbf{\Lambda}}}-\boldsymbol{M}_{\mathbf{\Lambda}^0\dot{\boldsymbol{H}}})(\mathbf{\Lambda}^0-\dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+)f_t^0. \tag{C.53}$$

Applying Lemmas C.3(i) and (v), we can find that $\frac{1}{N}X_t'\boldsymbol{M}_{\mathbf{\Lambda}^0\dot{\boldsymbol{H}}}(\mathbf{\Lambda}^0 - \dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+)f_t^0$ is the leading term, which will be the major focus in the following proof. Note that

$$\mathbf{\Lambda}^0 - \dot{\mathbf{\Lambda}}\dot{\boldsymbol{H}}^+ = \left(\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\dot{\mathbf{V}}_{NT} - \dot{\mathbf{\Lambda}}\dot{\mathbf{V}}_{NT}\right)\dot{\mathbf{V}}_{NT}^+\dot{\boldsymbol{H}}^+.$$

We can apply the decomposition (C.13) for $\mathbf{\Lambda}^0\dot{\boldsymbol{H}}\dot{\mathbf{V}}_{NT}-\dot{\mathbf{\Lambda}}\dot{\mathbf{V}}_{NT}$, use the fact that $\boldsymbol{M}_{\mathbf{\Lambda}^0\dot{\boldsymbol{H}}}\mathbf{\Lambda}^0\dot{\boldsymbol{H}} = \mathbf{0}$ and both $\dot{\boldsymbol{H}}$ and $\dot{\mathbf{V}}_{NT}$ are asymptotically invertible, and then obtain

$$\frac{1}{N}X_t'\boldsymbol{M}_{\mathbf{\Lambda}^0\dot{\boldsymbol{H}}}(\mathbf{\Lambda}^0 - \dot{\mathbf{\Lambda}}\boldsymbol{H}^+)f_t^0 = -\frac{1}{N}X_t'\boldsymbol{M}_{\mathbf{\Lambda}^0\dot{\boldsymbol{H}}}\left(\sum_{j=1}^{3}\dot{u}_{NT,j}+\sum_{j=6}^{8}\dot{u}_{NT,j}\right)\dot{\mathbf{V}}_{NT}^+\dot{\boldsymbol{H}}^+f_t^0. \tag{C.54}$$

Similar to the proof of Lemma C.3(i) and using the decomposition $\dot{\mathbf{\Lambda}} = (\dot{\mathbf{\Lambda}} - \mathbf{\Lambda}^0 \dot{\boldsymbol{H}}) + \mathbf{\Lambda}^0 \dot{\boldsymbol{H}}$, we may prove that

$$\frac{1}{N} \left\| X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} \left(\dot{u}_{NT,1} + \dot{u}_{NT,3} + \sum_{j=6}^{8} \dot{u}_{NT,j} \right) \dot{\mathbf{V}}_{NT}^{+} \dot{\boldsymbol{H}}^{+} f_t^0 \right\| = O_P\big(\delta_{NT}^{-1} + \dot{\eta}_{NT}\big). \tag{C.55}$$

Meanwhile, letting $\chi_{st} = f_s^{0\prime} \big(\frac{1}{T} \boldsymbol{F}^{0\prime} \boldsymbol{F}^0 \big)^{+} f_t^0$, we may also obtain

$$\begin{aligned} -\frac{1}{N} X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} \dot{u}_{NT,2} \dot{\mathbf{V}}_{NT}^{+} \dot{\boldsymbol{H}}^{+} f_t^0 &= \frac{1}{N^2 T} \sum_{s=1}^{T} X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_s d_s f_s^{0\prime} \mathbf{\Lambda}^{0\prime} \dot{\mathbf{\Lambda}} \dot{\mathbf{V}}_{NT}^{+} \dot{\boldsymbol{H}}^{+} f_t^0 \\ &= \frac{1}{NT} \sum_{s=1}^{T} X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_s \chi_{st} d_s. \end{aligned} \tag{C.56}$$

Note that

$$\frac{1}{N} X_t' \boldsymbol{M}_{\dot{\mathbf{\Lambda}}} X_t (\dot{\beta}_t - \beta_t^0) \overset{P}{\sim} \frac{1}{N} X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_t d_t, \tag{C.57}$$

where $a \overset{P}{\sim} b$ denotes $a = b(1 + o_P(1))$. By (C.39), (C.44), and (C.52)–(C.57), we have

$$\left\| \frac{1}{N} X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_t d_t - \frac{1}{NT} \sum_{s=1}^{T} X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_s \chi_{st} d_s \right\| = O_P \left(p^{1/2} N^{-1/2} + T^{-1/2} + \dot{\eta}_{NT} \right). \tag{C.58}$$

Let $\boldsymbol{L}_{NT} = \mathsf{diag} \left\{ \frac{1}{N} X_1' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_1, \ldots, \frac{1}{N} X_T' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_T \right\}$ and $\boldsymbol{L}_{NT,*}$ be the $T \times T$ block matrix with the (t, s) block being $\frac{1}{NT} X_t' \boldsymbol{M}_{\mathbf{\Lambda}^0 \dot{\boldsymbol{H}}} X_s \chi_{st}$. By (C.58), we may show that

$$\big(\boldsymbol{L}_{NT} - \boldsymbol{L}_{NT,*} \big) \dot{\boldsymbol{d}}_\beta = \boldsymbol{R}_{NT}, \tag{C.59}$$

where $\dot{\boldsymbol{d}}_\beta$ is defined in the proof of Lemma C.2, $\boldsymbol{R}_{NT} = (R_1', \ldots, R_T')'$ with

$$\|R_t\| = O_P \left(p^{1/2} N^{-1/2} + T^{-1/2} + \dot{\eta}_{NT} \right) \quad \text{and} \quad \frac{1}{T} \sum_{t=1}^{T} ||R_t||^2 = O_P \left(p N^{-1} + T^{-1} + \dot{\eta}_{NT}^2 \right).$$

Using the arguments as used in the proofs of Theorem 3.1 and Lemma C.2, we can prove that $\boldsymbol{L}_{NT} - \boldsymbol{L}_{NT,*}$ is asymptotically positive definite with the smallest eigenvalue bounded away from zero. Hence, (C.59) indicates that

$$\frac{1}{T} \|\dot{\boldsymbol{d}}_\beta\|^2 = \frac{1}{T} \sum_{t=1}^{T} \|\dot{\beta}_t - \beta_t^0\|^2 = O_P \left(p N^{-1} + T^{-1} + \dot{\eta}_{NT}^2 \right), \tag{C.60}$$

which, in conjunction with the definition of $\dot{\eta}_{NT}$ in the statement of Lemma C.3, implies that $\frac{1}{T} \|\dot{\boldsymbol{d}}_\beta\|^2 = O_P \left(p N^{-1} + T^{-1} \right)$, and strengthens the consistency result in Lemma C.2. By the fact that the matrix $\frac{1}{N}X_t'\boldsymbol{M}_{\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}}X_t$ is positive definite as well as (C.58) and (C.60), we can prove that

$$\left\|\dot{\beta}_t-\beta_t^0\right\|=O_P\left(p^{1/2}N^{-1/2}+T^{-1/2}\right)=O_P\left(\delta_{p,NT}^{-1}\right)$$

for each t, completing the proof of Lemma B.1 in Appendix B. ■

Proof of Lemma B.2. **(i)** Using the argument in the proof of Lemma C.2 (with some modifications), we may prove that $\eta_{NT}=o_P(1)$. Then, following the proofs of (C.44) and (C.52) above, we can readily show that

$$\frac{1}{N^2T}\sum_{t=1}^T\left\|X_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\varepsilon_t\right\|^2=O_P\big(pN^{-1}+T^{-1}\big). \tag{C.61}$$

Furthermore, by the Cauchy-Schwarz inequality, we have

$$\frac{1}{NT}\sum_{t=1}^T(\hat{\beta}_t-\beta_t^0)'X_t'\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\varepsilon_t=O_P\big(p^{1/2}\delta_{NT}^{-1}\big)\cdot\left(\frac{1}{T}\sum_{t=1}^T\left\|\hat{\beta}_t-\beta_t^0\right\|^2\right)^{1/2}=O_P\left(\delta_{p,NT}^{-1}\eta_{NT}^{1/2}\right). \tag{C.62}$$

(ii) As $\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\boldsymbol{\Lambda}^0}=\boldsymbol{0}$, we have $\sum_{t=1}^T f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}\varepsilon_t=\sum_{t=1}^T f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}(\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}-\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\varepsilon_t$. Similar to the decomposition in (C.37), we have

$$\boldsymbol{P}_{\hat{\boldsymbol{\Lambda}}}-\boldsymbol{P}_{\boldsymbol{\Lambda}^0\boldsymbol{H}}=\hat{\boldsymbol{\Lambda}}\big(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}\big)^+\hat{\boldsymbol{\Lambda}}'-\boldsymbol{\Lambda}^0\boldsymbol{H}\big(\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\boldsymbol{H}\big)^+\boldsymbol{H}'\boldsymbol{\Lambda}^{0\prime}\equiv\sum_{j=1}^7 v_{NT,j}, \tag{C.63}$$

where $\boldsymbol{H}\equiv\boldsymbol{H}_{NT}=\big(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\big)\big(\frac{1}{T}\boldsymbol{\Lambda}^{0\prime}\hat{\boldsymbol{\Lambda}}\big)\boldsymbol{V}_{NT}^+$, $\boldsymbol{V}_{NT}$ is defined in (2.7), and $v_{NT,j}$, $j=1,...,7$, are analogously defined as $\dot{v}_{NT,j}$ in the proof of Lemma C.3(v) with $\dot{\boldsymbol{\Lambda}}$ and $\dot{\boldsymbol{H}}$ replaced by $\hat{\boldsymbol{\Lambda}}$ and $\boldsymbol{H}$, respectively. We only need to show that

$$\left|\sum_{t=1}^T f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}(\boldsymbol{M}_{\hat{\boldsymbol{\Lambda}}}-\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\varepsilon_t\right|=\left|\sum_{t=1}^T f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\Big(\sum_{j=1}^7 v_{NT,j}\Big)\varepsilon_t\right|=O_P\left(\delta_{NT}^{-2}+\delta_{p,NT}^{-1}\eta_{NT}^{1/2}\right). \tag{C.64}$$

When $(\dot{\boldsymbol{\Lambda}},\dot{\boldsymbol{H}})$ is replaced by $(\hat{\boldsymbol{\Lambda}},\boldsymbol{H})$, it is easy to verify that the convergence results in Lemma C.3 still hold with $\dot{\eta}_{NT}$ replaced by η_{NT}. By Assumption 1(iii),

$$\left\|\sum_{t=1}^T\boldsymbol{\Lambda}^{0\prime}\varepsilon_t f_t^0\right\|=O_P(\sqrt{NT}), \tag{C.65}$$

which together with Lemma C.3 (with some modifications to allow the replacement of $\dot{\eta}_{NT}$, $\dot{\boldsymbol{\Lambda}}$, and $\dot{\boldsymbol{H}}$ by η_{NT}, $\hat{\boldsymbol{\Lambda}}$, and $\boldsymbol{H}$, respectively) indicates that

$$\frac{1}{NT}\left\|\sum_{t=1}^T f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}\left(v_{NT,2}+v_{NT,4}+v_{NT,7}\right)\varepsilon_t\right\|=O_P\big((NT)^{-1/2}(\delta_{NT}^{-2}+\eta_{NT}^{1/2})\big). \tag{C.66}$$

On the other hand, note that

$$\left\|\sum_{t=1}^{T}\left(\hat{\boldsymbol{\Lambda}}\boldsymbol{V}_{NT}-\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{V}_{NT}\right)'\varepsilon_t f_t^0\right\|=\left\|\sum_{t=1}^{T}\left(\sum_{j=1}^{8}u_{NT,j}\right)'\varepsilon_t f_t^0\right\|, \qquad \text{(C.67)}$$

where $u_{NT,j}$, $j=1,...,8$, are defined similarly to $\dot{u}_{NT,j}$ in the proof of Lemma C.3 (i) with $\dot{\beta}_t$ and $\dot{\boldsymbol{\Lambda}}$ replaced by $\hat{\beta}_t$ and $\hat{\boldsymbol{\Lambda}}$, respectively. Let $\hat{d}_s=\hat{\beta}_s-\beta_s^0$. Then, by the definition of $u_{NT,j}$ and using Assumptions 1(i)–(iii), we can prove that

$$\begin{aligned}\left\|\sum_{t=1}^{T}u_{NT,1}'\varepsilon_t f_t^0\right\| &= \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{\boldsymbol{\Lambda}}'X_s\hat{d}_s\hat{d}_s'X_s'\varepsilon_t f_t^0\right\| \\ &= O_P\big(T^{-1}\big)\cdot\sum_{t=1}^{T}\|f_t^0\|\sum_{s=1}^{T}\|\hat{d}_s\|^2\|X_s'\varepsilon_t\| \\ &= O_P\left(N^{1/2}Tp^{1/2}\eta_{NT}\right), \qquad \text{(C.68)}\end{aligned}$$

and

$$\begin{aligned}\left\|\sum_{t=1}^{T}u_{NT,2}'\varepsilon_t f_t^0\right\| &= \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{\boldsymbol{\Lambda}}'\boldsymbol{\Lambda}^0 f_s^0\hat{d}_s'X_s'\varepsilon_t f_t^0\right\| \\ &= O_P(T^{-1})\sum_{t=1}^{T}\|f_t^0\|\sum_{s=1}^{T}\|\hat{d}_s\|\|f_s^0\|\|X_s'\varepsilon_t\| \\ &= O_P\big(N^{1/2}T(p\eta_{NT})^{1/2}\big). \qquad \text{(C.69)}\end{aligned}$$

By analogous arguments, we can also show that

$$\left\|\sum_{t=1}^{T}u_{NT,4}'\varepsilon_t f_t^0\right\|=O_P\big(N^{1/2}T\eta_{NT}^{1/2}\big). \qquad \text{(C.70)}$$

On the other hand, using Lemma C.3 we can show that

$$\begin{aligned}\left\|\sum_{t=1}^{T}u_{NT,3}'\varepsilon_t f_t^0\right\| &= \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{\boldsymbol{\Lambda}}'\varepsilon_s\hat{d}_s'X_s'\varepsilon_t f_t^0\right\| \\ &\leq \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{H}'\boldsymbol{\Lambda}_0'\varepsilon_s\hat{d}_s'X_s'\varepsilon_t f_t^0\right\|+\frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}_0\boldsymbol{H}\right)'\varepsilon_s\hat{d}_s'X_s'\varepsilon_t f_t^0\right\| \\ &\leq \|\boldsymbol{H}\|\left(\frac{1}{NT}\sum_{s=1}^{T}||\boldsymbol{\Lambda}_0'\varepsilon_s||^2\right)^{1/2}\left(\frac{1}{NT}\sum_{s=1}^{T}\left\|\hat{d}_s'\sum_{s=1}^{T}X_s'\varepsilon_t f_t^0\right\|^2\right)^{1/2} \\ &\quad +\left(\frac{1}{NT}\sum_{s=1}^{T}||(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}_0\boldsymbol{H})'\varepsilon_s||^2\right)^{1/2}\left(\frac{1}{NT}\sum_{s=1}^{T}\left\|\hat{d}_s'\sum_{t=1}^{T}X_s'\varepsilon_t f_t^0\right\|^2\right)^{1/2}\end{aligned}$$

$$
\begin{aligned}
&= O_P\left(T(p\eta_{NT})^{1/2}\right) + O_P\left((1+N^{1/2}T^{-1/2})(\delta_{NT}^{-1}+\eta_{NT}^{1/2})T(p\eta_{NT})^{1/2}\right) \\
&= O_P\left((1+N^{1/2}T^{-1/2}\delta_{NT}^{-1}+N^{1/2}T^{-1/2}\eta_{NT}^{1/2})T(p\eta_{NT})^{1/2}\right), \qquad \text{(C.71)}
\end{aligned}
$$

and analogously

$$
\begin{aligned}
\left\|\sum_{t=1}^{T} u_{NT,5}'\varepsilon_t f_t^0\right\| &= \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{\boldsymbol{\Lambda}}'\varepsilon_s f_s^{0\prime}\boldsymbol{\Lambda}^{0\prime}\varepsilon_t f_t^0\right\| \\
&\leq \frac{1}{NT}\left\|\sum_{s=1}^{T}\sum_{t=1}^{T}\boldsymbol{H}'\boldsymbol{\Lambda}_0'\varepsilon_s f_s^{0\prime}\boldsymbol{\Lambda}^{0\prime}\varepsilon_t f_t^0\right\| + \frac{1}{NT}\left\|\sum_{s=1}^{T}\sum_{t=1}^{T}\left(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}_0\boldsymbol{H}\right)'\varepsilon_s f_s^{0\prime}\boldsymbol{\Lambda}^{0\prime}\varepsilon_t f_t^0\right\| \\
&\leq \|\boldsymbol{H}\|\frac{1}{NT}\|\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0\|^2 + \frac{1}{NT}\left\|\sum_{s=1}^{T}\left(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}_0\boldsymbol{H}\right)'\varepsilon_s f_s^{0\prime}\right\|\left\|\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\varepsilon}\boldsymbol{F}^0\right\| \\
&= O_P(1) + O_P\left(N^{1/2}T^{1/2}(\delta_{NT}^{-2}+\eta_{NT}^{1/2})\right). \qquad \text{(C.72)}
\end{aligned}
$$

Using the fact that under Assumptions 1(i) and (iv)

$$
\sum_{s=1}^{T}\left\|\sum_{t=1}^{T}\varepsilon_s'\varepsilon_t f_t^0\right\|^2 \leq \left(\sum_{s=1}^{T}\sum_{t_1=1}^{T}\|\varepsilon_s'\varepsilon_{t_1}\|^2\right)\left(\sum_{t_2=1}^{T}\|f_{t_2}^0\|^2\right) = O_P\left(T^2N(N+T)\right), \qquad \text{(C.73)}
$$

we have

$$
\begin{aligned}
\left\|\sum_{t=1}^{T} u_{NT,6}'\varepsilon_t f_t^0\right\| &= \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{\boldsymbol{\Lambda}}'X_s\hat{d}_s\varepsilon_s'\varepsilon_t f_t^0\right\| \\
&\leq \frac{1}{NT}\max_{1\leq s\leq T}\|\hat{\boldsymbol{\Lambda}}'X_s\|\cdot\left(\sum_{s=1}^{T}\|\hat{d}_s\|^2\right)^{1/2}\cdot\left(\sum_{s=1}^{T}\left\|\sum_{t=1}^{T}\varepsilon_s'\varepsilon_t f_t^0\right\|^2\right)^{1/2} \\
&= O_P(T^{-1})\cdot O_P\left(T^{1/2}\eta_{NT}^{1/2}\right)\cdot O_P\left(TN^{1/2}(N^{1/2}+T^{1/2})\right) \\
&= O_P\left(\eta_{NT}^{1/2}(NT^{1/2}+N^{1/2}T)\right). \qquad \text{(C.74)}
\end{aligned}
$$

Notice that

$$
\begin{aligned}
\left\|\sum_{t=1}^{T} u_{NT,8}'\varepsilon_t f_t^0\right\| &= \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{\boldsymbol{\Lambda}}'\varepsilon_s\varepsilon_s'\varepsilon_t f_t^0\right\| \\
&\leq \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{H}'\boldsymbol{\Lambda}_0'\varepsilon_s\varepsilon_s'\varepsilon_t f_t^0\right\| + \frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}_0\boldsymbol{H}\right)'\varepsilon_s\varepsilon_s'\varepsilon_t f_t^0\right\|.
\end{aligned}
$$

For the first term on the right hand side, by the Cauchy-Schwarz inequality and Assumption 1(iii) and (C.73) we may show that

$$
\begin{aligned}
\frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{H}'\boldsymbol{\Lambda}_0'\varepsilon_s\varepsilon_s'\varepsilon_t f_t^0\right\| &\leq \frac{1}{NT}\|\boldsymbol{H}\|\cdot\left(\sum_{s=1}^{T}\|\boldsymbol{\Lambda}_0'\varepsilon_s\|^2\right)^{1/2}\cdot\left(\sum_{s=1}^{T}\|\sum_{t=1}^{T}\varepsilon_s'\varepsilon_t f_t^0\|^2\right)^{1/2} \\
&= O_P\big((NT)^{-1/2}\big)O_P\big(TN^{1/2}(N^{1/2}+T^{1/2})\big) = O_P\big((NT)^{1/2}+T\big).
\end{aligned}
$$

For the second term on the right hand side, by Lemma C.3(vii) (with $\dot{\eta}_{NT}$, $\dot{\boldsymbol{\Lambda}}$, and $\dot{\boldsymbol{H}}$ replaced by η_{NT}, $\hat{\boldsymbol{\Lambda}}$, and $\boldsymbol{H}$, respectively), we have

$$
\begin{aligned}
\frac{1}{NT}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}_0\boldsymbol{H}\right)'\varepsilon_s\varepsilon_s'\varepsilon_t f_t^0\right\| &\leq \left(\frac{1}{NT}\sum_{s=1}^{T}\|\left(\hat{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}_0\boldsymbol{H}\right)'\varepsilon_s\|^2\right)^{1/2}\cdot\left(\frac{1}{NT}\sum_{s=1}^{T}\|\sum_{t=1}^{T}\varepsilon_s'\varepsilon_t f_t^0\|^2\right)^{1/2} \\
&= O_P\left((1+N^{1/2}T^{-1/2})(\delta_{NT}^{-1}+\eta_{NT}^{1/2})\right)O_P\left(T+N^{1/2}T^{1/2}\right) \\
&= O_P\left((T+N)(\delta_{NT}^{-1}+\eta_{NT}^{1/2})\right).
\end{aligned}
$$

It follows that

$$
\left\|\sum_{t=1}^{T}u_{NT,8}'\varepsilon_t f_t^0\right\| = O_P\left((NT)^{1/2}+T+N\eta_{NT}^{1/2}\right). \tag{C.75}
$$

Finally, noting that $\left|\sum_{s=1}^{T}\sum_{t=1}^{T}f_s^{0\prime}\varepsilon_s'\varepsilon_t f_t^0\right| = O_P(NT)$ by Assumption 1(iv), we can also show that

$$
\left\|\sum_{t=1}^{T}u_{NT,7}'\varepsilon_t f_t^0\right\| = O_P(N). \tag{C.76}
$$

By (C.67)–(C.76), we have

$$
\frac{1}{NT}\left\|\sum_{t=1}^{T}\left(\hat{\boldsymbol{\Lambda}}\boldsymbol{V}_{NT}-\boldsymbol{\Lambda}^0\boldsymbol{H}\boldsymbol{V}_{NT}\right)'\varepsilon_t f_t^0\right\| = O_P\left(\delta_{NT}^{-2}+\delta_{p,NT}^{-1}\eta_{NT}^{1/2}\right). \tag{C.77}
$$

With this, we readily prove that

$$
\frac{1}{NT}\left\|\sum_{t=1}^{T}f_t^{0\prime}\boldsymbol{\Lambda}^{0\prime}(v_{NT,1}+v_{NT,3}+v_{NT,5}+v_{NT,6})\varepsilon_t\right\| = O_P\left(\delta_{NT}^{-2}+\delta_{p,NT}^{-1}\eta_{NT}^{1/2}\right), \tag{C.78}
$$

which together with (C.66), leads to (C.64). Hence, we complete the proof of (ii).

(iii) This follows from Lemmas C.1(iii) and (iv). ■

Before proving Lemma B.3 in Appendix B, we need to introduce two technical lemmas. The first lemma is similar to Lemma C.3 with the preliminary estimates replaced by the post-LASSO estimates. Let $\tilde{\boldsymbol{\Lambda}}_{m^0} = \tilde{\boldsymbol{\Lambda}}(\mathcal{T}_{m^0}^0)$ be the infeasible estimate of the factor loadings in the post-LASSO estimation procedure, $\tilde{\boldsymbol{H}} = \left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\,\tilde{\boldsymbol{\Lambda}}_{m^0}\right)\tilde{\boldsymbol{V}}_{NT}^{+}$ with $\tilde{\boldsymbol{V}}_{NT}$ defined in the proof of Theorem 3.4 in Appendix B, and $\tilde{\eta}_{NT} = \frac{1}{m^0}\sum_{j=1}^{m^0+1}\|\tilde{\alpha}_{m^0j}-\alpha_j^0\|^2$, where $\tilde{\alpha}_{m^0j}$ is the j-th p-dimensional element of the infeasible estimate $\tilde{\boldsymbol{\alpha}}_{m^0} = \tilde{\boldsymbol{\alpha}}_{m^0}(\mathcal{T}_{m^0}^0)$.

Lemma C.4 *Suppose that the conditions in Theorem 3.4 hold. Then we have*

(i) $\frac{1}{N}\left\|\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right\|^2 = O_P\left(\delta_{NT}^{-2}+\tilde{\eta}_{NT}\right)$,

(ii) $\frac{1}{N}\big(\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\big)'\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}=O_P\big(\delta_{NT}^{-2}+\tilde{\eta}_{NT}^{1/2}\big)$,
(iii) $\frac{1}{N}\big(\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\big)'\tilde{\mathbf{\Lambda}}_{m^0}=O_P\big(\delta_{NT}^{-2}+\tilde{\eta}_{NT}^{1/2}\big)$,
(iv) $\frac{1}{N}\big(\tilde{\mathbf{\Lambda}}_{m^0}'\tilde{\mathbf{\Lambda}}_{m^0}-\tilde{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\big)=O_P\big(\delta_{NT}^{-2}+\tilde{\eta}_{NT}^{1/2}\big)$,
(v) $\big\|\boldsymbol{P}_{\tilde{\mathbf{\Lambda}}_{m^0}}-\boldsymbol{P}_{\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}}\big\|=O_P\big(\delta_{NT}^{-1}+\tilde{\eta}_{NT}^{1/2}\big)$,
(vi) $\frac{1}{NT}\sum_{s=1}^{T}(\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}})'\varepsilon_s\gamma_s'=O_P\big(\delta_{NT}^{-2}+\tilde{\eta}_{NT}^{1/2}\big)$ *with* $\gamma_s=1$ *or* f_s^0, *and*
(vii) $\frac{1}{NT}\sum_{s=1}^{T}||(\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}})'\varepsilon_s||^2=O_P\big((1+NT^{-1})(\delta_{NT}^{-2}+\tilde{\eta}_{NT})\big)$.

Proof of Lemma C.4. The proof is analogous to that of Lemma C.3. Hence, we only sketch it. For notational simplicity, we let $\tilde{\boldsymbol{V}}\equiv\tilde{\boldsymbol{V}}_{NT}$, and $\tilde{\eta}_j=\tilde{\alpha}_{m^0j}-\alpha_j^0$, $j=1,...,m^0+1$. By (B.25) in the proof of Theorem 3.4, we have

$$
\begin{aligned}
&\tilde{\mathbf{\Lambda}}_{m^0}\tilde{\boldsymbol{V}}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\tilde{\boldsymbol{V}}\\
=&\left[\frac{1}{NT}\sum_{j=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}\big(Y_t-X_t\tilde{\alpha}_{m^0j}\big)\big(Y_t-X_t\tilde{\alpha}_{m^0j}\big)'\right]\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\tilde{\boldsymbol{V}}\\
=&\left[\frac{1}{NT}\sum_{j=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}\big(-X_t\tilde{\eta}_j+\mathbf{\Lambda}^0f_t^0+\varepsilon_t\big)\big(-X_t\tilde{\eta}_j+\mathbf{\Lambda}^0f_t^0+\varepsilon_t\big)'\right]\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\tilde{\boldsymbol{V}}\\
=&\frac{1}{NT}\sum_{j=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t\tilde{\eta}_j\tilde{\eta}_j'X_t'\tilde{\mathbf{\Lambda}}_{m^0}-\frac{1}{NT}\sum_{j=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t\tilde{\eta}_jf_t^{0\prime}\mathbf{\Lambda}^{0\prime}\tilde{\mathbf{\Lambda}}_{m^0}-\frac{1}{NT}\sum_{j=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t\tilde{\eta}_j\varepsilon_t'\tilde{\mathbf{\Lambda}}_{m^0}\\
&-\frac{1}{NT}\sum_{j=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}\mathbf{\Lambda}^0f_t^0\tilde{\eta}_j'X_t'\tilde{\mathbf{\Lambda}}_{m^0}+\frac{1}{NT}\sum_{t=1}^{T}\mathbf{\Lambda}^0f_t^0\varepsilon_t'\tilde{\mathbf{\Lambda}}_{m^0}-\frac{1}{NT}\sum_{j=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}\varepsilon_t\tilde{\eta}_j'X_t'\tilde{\mathbf{\Lambda}}_{m^0}\\
&+\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_tf_t^{0\prime}\mathbf{\Lambda}^{0\prime}\tilde{\mathbf{\Lambda}}_{m^0}+\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t\varepsilon_t'\tilde{\mathbf{\Lambda}}_{m^0}\\
\equiv&\sum_{j=1}^{8}\tilde{u}_{NT,j}. \qquad \text{(C.79)}
\end{aligned}
$$

Then following the proof of Lemma C.3 with $\dot{\mathbf{\Lambda}}$ and d_t replaced by $\tilde{\mathbf{\Lambda}}_{m^0}$ and $\tilde{\eta}_j$, respectively, and using Assumption 3(ii), we can readily prove Lemma C.4(i). Note that

$$\frac{1}{N}\big(\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\big)'\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}=\frac{1}{N}\sum_{j=1}^{8}\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,j}'\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\equiv\frac{1}{N}\sum_{j=1}^{8}\tilde{u}_{NT,j}^{*}. \qquad \text{(C.80)}$$

Then following the proof of Lemma C.3(ii) and using Lemma C.4(i), we readily prove Lemma C.4(ii). The results in (iii) and (iv) can be proved by combining Lemmas C.4(i) and (ii). Similar to (C.37), we have the following decomposition:

$$\boldsymbol{P}_{\tilde{\boldsymbol{\Lambda}}_{m^0}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}} = \tilde{\boldsymbol{\Lambda}}_{m^0}\big(\tilde{\boldsymbol{\Lambda}}'_{m^0}\tilde{\boldsymbol{\Lambda}}_{m^0}\big)^+\tilde{\boldsymbol{\Lambda}}'_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime} \equiv \sum_{j=1}^{7}\tilde{v}_{NT,j}, \qquad \text{(C.81)}$$

where

$$\begin{aligned}
\tilde{v}_{NT,1} &= \big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)\big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)',\\
\tilde{v}_{NT,2} &= \big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)\big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime},\\
\tilde{v}_{NT,3} &= \big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)\big[\big(\tilde{\boldsymbol{\Lambda}}'_{m^0}\tilde{\boldsymbol{\Lambda}}_{m^0}\big)^+ - \big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\big]\big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)',\\
\tilde{v}_{NT,4} &= \big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)\big[\big(\tilde{\boldsymbol{\Lambda}}'_{m^0}\tilde{\boldsymbol{\Lambda}}_{m^0}\big)^+ - \big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\big]\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime},\\
\tilde{v}_{NT,5} &= \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)',\\
\tilde{v}_{NT,6} &= \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big[\big(\tilde{\boldsymbol{\Lambda}}'_{m^0}\tilde{\boldsymbol{\Lambda}}_{m^0}\big)^+ - \big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\big]\big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)',\\
\tilde{v}_{NT,7} &= \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big[\big(\tilde{\boldsymbol{\Lambda}}'_{m^0}\tilde{\boldsymbol{\Lambda}}_{m^0}\big)^+ - \big(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^+\big]\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}.
\end{aligned}$$

By (C.81) and Lemmas C.4(i) and (iv), we can prove (v). The proofs of (vi) and (vii) parallel to those of Lemmas C.3(vi) and (vii). We have thus completed the proof of Lemma C.4. ■

Lemma C.5 *Suppose that the conditions in Theorem 3.4 hold. Then we have*

(i) $\tilde{\eta}_{NT} = \frac{1}{m^0}\sum_{j=1}^{m^0+1}\|\tilde{\alpha}_{m^0j} - \alpha_j^0\|^2 = O_P\big(\delta_{p,NT}^{-2}\big)$,

(ii) $\frac{1}{N}\big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)'\varepsilon_t = \tilde{\boldsymbol{H}}'\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+\left(\frac{1}{NT}\sum_{s=1}^T f_s^0\varepsilon_s'\varepsilon_t\right) + O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0} - \boldsymbol{\alpha}^0\|\right)$ $+O_P\left(\delta_{p,NT}^{-3}\right)$ *for* $t = 1, ..., T$,

(iii) $\frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\left(\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^+\left(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)'\varepsilon_t - \sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{\Lambda}^0\left(\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\right)^+ \left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+\left(\frac{1}{NT}\sum_{s=1}^T f_s^0\varepsilon_s'\varepsilon_t\right)\right\| = O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0} - \boldsymbol{\alpha}^0\|\right) + O_P\left(\delta_{p,NT}^{-3}\right)$ *for* $j = 1,\ldots,m^0+1$,

(iv) $\frac{1}{NT}\sum_{t=1}^T\left\|(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}})'\varepsilon_t f_t^0\right\| = O_P\left(\delta_{p,NT}^{-2}\right)$.

Proof of Lemma C.5. As the proof of the convergence rates for $\tilde{\boldsymbol{\alpha}}_{m^0}$ in (i) is similar to the proof of Lemma B.1, we omit the details. Furthermore, the results in (iii) and (iv) can be easily proved by using (ii). Hence we only focus on the proof of the result in (ii).

Note that for any $t = 1, ..., T$,

$$\frac{1}{N}\big(\tilde{\boldsymbol{\Lambda}}_{m^0} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)'\varepsilon_t = \frac{1}{N}\tilde{\boldsymbol{V}}^+\big(\tilde{\boldsymbol{\Lambda}}_{m^0}\tilde{\boldsymbol{V}} - \boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\tilde{\boldsymbol{V}}\big)'\varepsilon_t = \frac{1}{N}\tilde{\boldsymbol{V}}^+\big(\sum_{k=1}^{8}\tilde{u}_{NT,k}\big)'\varepsilon_t \qquad \text{(C.82)}$$

by using (C.79) in the proof of Lemma C.4. By Lemma C.5(i), Assumptions 1(ii), (iii) and 3(ii), and the Jensen inequality, we have

$$
\begin{aligned}
\frac{1}{N}\left\|\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,1}'\varepsilon_t\right\| &= \frac{1}{N^2T}\left\|\tilde{\boldsymbol{V}}^{+}\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\tilde{\boldsymbol{\Lambda}}_{m^0}'X_s\tilde{\eta}_k\tilde{\eta}_k'X_s'\varepsilon_t\right\| \\
&= O_P\big(N^{-2}T^{-1}\big)\left\|\tilde{\boldsymbol{\Lambda}}_{m^0}\right\|\max_{1\le s\le T}\mu_{\max}^{1/2}(X_s'X_s)\cdot\sum_{k=1}^{m^0+1}\|\tilde{\eta}_k\|^2\sum_{s=T_{k-1}^0}^{T_k^0-1}\left\|X_s'\varepsilon_t\right\| \\
&= O_P\left(p^{1/2}N^{-1/2}\tilde{\eta}_{NT}\right) = O_P\left(\delta_{p,NT}^{-3}\right). \qquad \text{(C.83)}
\end{aligned}
$$

By Lemmas C.4(i) and C.5(i) and Assumptions 1(iii), (iv) and 3(ii), we can show that

$$
\begin{aligned}
&\frac{1}{N}\|\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,3}'\varepsilon_t\| \\
=\ & \frac{1}{N^2T}\left\|\tilde{\boldsymbol{V}}^{+}\left[\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\tilde{\boldsymbol{H}}'\boldsymbol{\Lambda}^{0\prime}\varepsilon_s\tilde{\eta}_k'X_s'\varepsilon_t + \sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\big(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)'\varepsilon_s\tilde{\eta}_k'X_s'\varepsilon_t\right]\right\| \\
=\ & O_P\big(N^{-2}T^{-1}\big)\left[\sum_{k=1}^{m^0+1}\|\tilde{\eta}_k\|\sum_{s=T_{k-1}^0}^{T_k^0-1}\|\boldsymbol{\Lambda}^{0\prime}\varepsilon_s\|\left\|X_s'\varepsilon_t\right\| + \|\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\|\sum_{k=1}^{m^0+1}\|\tilde{\eta}_k\|\sum_{s=T_{k-1}^0}^{T_k^0-1}\|\varepsilon_s\|\left\|X_s'\varepsilon_t\right\|\right] \\
=\ & O_P\left(N^{-1}(p\tilde{\eta}_{NT})^{1/2}\right) + O_P\left(N^{-1/2}(\delta_{NT}^{-1}+\tilde{\eta}_{NT}^{1/2})(p\tilde{\eta}_{NT})^{1/2}\right) \\
=\ & O_P\left(N^{-1}(p\tilde{\eta}_{NT})^{1/2}+\delta_{p,NT}^{-3}\right). \qquad \text{(C.84)}
\end{aligned}
$$

By Assumptions 1(i), (iii) and 3(ii), and Lemma C.5(i), we have

$$
\begin{aligned}
\frac{1}{N}\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,4}'\varepsilon_t &= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\left(\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\tilde{\boldsymbol{\Lambda}}_{m^0}'X_s\tilde{\eta}_kf_s^{0\prime}\boldsymbol{\Lambda}^{0\prime}\right)\varepsilon_t \\
&= O_P\big(N^{-2}T^{-1}\big)\cdot\sum_{k=1}^{m^0+1}\|\tilde{\eta}_k\|\left(\sum_{s=T_{k-1}^0}^{T_k^0-1}\|\tilde{\boldsymbol{\Lambda}}_{m^0}'X_s\|\|f_s^0\|\|\sum_{i=1}^{N}\lambda_i^0\varepsilon_{it}\|\right) \\
&= O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right). \qquad \text{(C.85)}
\end{aligned}
$$

Analogously, we can show that

$$
\frac{1}{N}\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,2}'\varepsilon_t = O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right). \qquad \text{(C.86)}
$$

By Assumptions 1(iii) and (iv), we can prove that

$$
\begin{aligned}
\frac{1}{N}\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,5}^{\prime}\varepsilon_t &= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\left(\sum_{s=1}^{T}\boldsymbol{\Lambda}^0 f_s^0\varepsilon_s^{\prime}\tilde{\boldsymbol{\Lambda}}_{m^0}\right)^{\prime}\varepsilon_t \\
&= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^{\prime}\varepsilon_s f_s^0\boldsymbol{\Lambda}^{0\prime}\varepsilon_t+\frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\left(\sum_{s=1}^{T}\tilde{\boldsymbol{H}}^{\prime}\boldsymbol{\Lambda}^{0\prime}\varepsilon_s f_s^{0\prime}\boldsymbol{\Lambda}^{0\prime}\varepsilon_t\right) \\
&= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^{\prime}\varepsilon_s f_s^0\boldsymbol{\Lambda}^{0\prime}\varepsilon_t+O_P\left(\frac{1}{N^2T}\Big\|\sum_{s=1}^{T}\boldsymbol{\Lambda}^{0\prime}\varepsilon_s f_s^0\Big\|\|\boldsymbol{\Lambda}^{0\prime}\varepsilon_t\|\right) \\
&= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^{\prime}\varepsilon_s f_s^0\boldsymbol{\Lambda}^{0\prime}\varepsilon_t+O_P\left(\delta_{NT}^{-3}\right). \qquad \text{(C.87)}
\end{aligned}
$$

By Assumptions 1(ii), (iv) and Lemma C.5(i), we have

$$
\begin{aligned}
\frac{1}{N}\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,6}^{\prime}\varepsilon_t &= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\left(\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\tilde{\boldsymbol{\Lambda}}_{m^0}^{\prime}X_s\tilde{\eta}_k\varepsilon_s^{\prime}\right)\varepsilon_t \\
&= O_P\big(N^{-2}T^{-1}\big)\cdot\sum_{k=1}^{m^0+1}\|\tilde{\eta}_k\|\left[\sum_{s=T_{k-1}^0}^{T_k^0-1}\|\tilde{\boldsymbol{\Lambda}}_{m^0}^{\prime}X_s\|\left\|\sum_{i=1}^{N}\varepsilon_{it}\varepsilon_{is}\right\|\right] \\
&= O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right) \qquad \text{(C.88)}
\end{aligned}
$$

By the definition of $\tilde{\boldsymbol{H}}$ and noting that $\tilde{\boldsymbol{V}}_{NT}^{+}$ is diagonal, we have

$$
\frac{1}{N}\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,7}^{\prime}\varepsilon_t=\left(\frac{1}{N}\tilde{\boldsymbol{V}}^{+}\tilde{\boldsymbol{\Lambda}}_{m^0}^{\prime}\boldsymbol{\Lambda}^0\right)\left[\frac{1}{NT}\sum_{s=1}^{T}f_s^0\varepsilon_s^{\prime}\varepsilon_t\right]=\tilde{\boldsymbol{H}}^{\prime}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{+}\left[\frac{1}{NT}\sum_{s=1}^{T}f_s^0\varepsilon_s^{\prime}\varepsilon_t\right]. \qquad \text{(C.89)}
$$

By the definition of $\tilde{u}_{NT,8}$ and Assumption 3(iii),

$$
\begin{aligned}
\frac{1}{N}\tilde{\boldsymbol{V}}^{+}\tilde{u}_{NT,8}^{\prime}\varepsilon_t &= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^{\prime}\varepsilon_s\varepsilon_s^{\prime}\varepsilon_t+\frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\tilde{\boldsymbol{H}}^{\prime}\sum_{s=1}^{T}\boldsymbol{\Lambda}^{0\prime}\varepsilon_s\varepsilon_s^{\prime}\varepsilon_t \\
&= \frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^{\prime}\varepsilon_s\varepsilon_s^{\prime}\varepsilon_t+O_P\left(\delta_{NT}^{-3}\right). \qquad \text{(C.90)}
\end{aligned}
$$

Combining the results in (C.82)–(C.90) yields

$$
\begin{aligned}
\frac{1}{N}\big(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\big)^{\prime}\varepsilon_t &= \tilde{\boldsymbol{H}}^{\prime}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{+}\frac{1}{NT}\sum_{s=1}^{T}f_s^0\varepsilon_s^{\prime}\varepsilon_t+\frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^{\prime}\varepsilon_s f_s^0\boldsymbol{\Lambda}^{0\prime}\varepsilon_t \\
&\quad+\frac{1}{N^2T}\tilde{\boldsymbol{V}}^{+}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{\Lambda}}_{m^0}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}}\right)^{\prime}\varepsilon_s\varepsilon_s^{\prime}\varepsilon_t+O_P\left(\delta_{p,NT}^{-3}\right) \\
&\quad+O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right). \qquad \text{(C.91)}
\end{aligned}
$$

By Assumptions 1(i) and (iv), the first term on the right hand side of (C.91) is $O_P(\delta_{NT}^{-2})$; by Assumptions 1(iii) and Lemmas C.4(vi) and C.5(i) we can show the second term is $O_P(\delta_{p,NT}^{-1}\delta_{NT}^{-1})$; by Assumptions 1(iii) and (iv) and Lemma C.4(vii) and , we can show the third and fourth terms are $O_P(\delta_{p,NT}^{-1}\delta_{NT}^{-1})$. It follows that

$$\frac{1}{N}\left(\tilde{\mathbf{\Lambda}}_{m^0}-\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\right)'\varepsilon_t = O_P\left(\delta_{p,NT}^{-1}\delta_{NT}^{-1}\right). \tag{C.92}$$

By (C.92) and following the above arguments, we can further show that the second and third terms on the right hand side of (C.91) are $O_P(\delta_{p,NT}^{-3})$. This completes the proof of Lemma C.5(ii). ■

Proof of Lemma B.3. For notional simplicity, we let $\tilde{\mathbf{\Lambda}}=\tilde{\mathbf{\Lambda}}_{m^0}$ throughout this proof.

(i) Noting that

$$-(\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}}-\boldsymbol{M}_{\mathbf{\Lambda}^0}) = \tilde{\mathbf{\Lambda}}\big(\tilde{\mathbf{\Lambda}}'\tilde{\mathbf{\Lambda}}\big)^{+}\tilde{\mathbf{\Lambda}}' - \mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\big(\tilde{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\big)^{+}\tilde{\boldsymbol{H}}'\mathbf{\Lambda}^{0\prime} = \sum_{k=1}^{7}\tilde{v}_{NT,k} \tag{C.93}$$

and by using the decomposition (C.81), we have

$$\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'(\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}}-\boldsymbol{M}_{\mathbf{\Lambda}^0})\varepsilon_t = -\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\left(\sum_{k=1}^{7}\tilde{v}_{NT,k}\right)\varepsilon_t. \tag{C.94}$$

By (C.94), Lemmas C.4(i), (iv) and C.5(iii), we can prove that for any $j=1,...,m^0+1$,

$$\begin{aligned}
&\left\|\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'(\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}}-\boldsymbol{M}_{\mathbf{\Lambda}^0})\varepsilon_t + B_{NT,j}(2,1)\right\| \\
\leq\ & \left\|\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\Big(\sum_{j=1,\neq 5}^{7}\tilde{v}_{NT,j}\Big)\varepsilon_t\right\| + \left\|\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\tilde{v}_{NT,5}\varepsilon_t - B_{NT,j}(2,1)\right\| \\
=\ & \left\|\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\tilde{v}_{NT,5}\varepsilon_t - B_{NT,j}(2,1)\right\| + O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m_0}-\boldsymbol{\alpha}^0\| + \delta_{p,NT}^{-3}\right) \\
=\ & O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m_0}-\boldsymbol{\alpha}^0\| + \delta_{p,NT}^{-3}\right)
\end{aligned} \tag{C.95}$$

which completes the proof of Lemma B.3(i).

(ii) Noting that for any $j=1,...,m^0+1$,

$$\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}}\big(\mathbf{\Lambda}^0-\tilde{\mathbf{\Lambda}}\tilde{\boldsymbol{H}}^{+}\big)f_t^0 = \frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}}\big(\mathbf{\Lambda}^0\tilde{\boldsymbol{H}}\tilde{\boldsymbol{V}}-\tilde{\mathbf{\Lambda}}\tilde{\boldsymbol{V}}\big)\tilde{\boldsymbol{V}}^{+}\tilde{\boldsymbol{H}}^{+}f_t^0,$$

and $\tilde{\boldsymbol{V}}^{+}\tilde{\boldsymbol{H}}^{+} = \left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^{0}\right)^{+}$, by the decomposition (C.79), we have

$$\begin{aligned}
&\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\left(\boldsymbol{\Lambda}^0 - \tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}^{+}\right)f_t^0 \\
= \;& -\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\left(\sum_{l=1}^{8}\tilde{u}_{NT,l}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^{0}\right)^{+} f_t^0. \qquad \text{(C.96)}
\end{aligned}$$

We next analyze each term on the right hand side of the equation (C.96).

For $l=1$, by the definition of $\tilde{u}_{NT,1}$, Assumptions 1(i)(ii), and Lemma C.5(i), we have

$$\begin{aligned}
&\frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,1}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^{0}\right)^{+} f_t^0\right\| \\
= \;& \frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\left(\frac{1}{NT}\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1} X_s\tilde{\eta}_k\tilde{\eta}_k'X_s'\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^{0}\right)^{+} f_t^0\right\| \\
= \;& O_P\left(\frac{1}{N^2T}\sum_{k=1}^{m^0+1}\|\tilde{\eta}_k\|^2\cdot\frac{1}{\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\|X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}X_s\|\|X_s'\tilde{\boldsymbol{\Lambda}}\|\|f_t^0\|\right) \\
= \;& O_P\left(p\tilde{\eta}_{NT}\right) = O_P\left(p\delta_{p,NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right). \qquad \text{(C.97)}
\end{aligned}$$

For $l=2$, by the definition of $\tilde{u}_{NT,2}$, we have

$$\begin{aligned}
&\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,2}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^{0}\right)^{+} f_t^0 \\
= \;& -\frac{1}{NT\tau_j(T)}\sum_{k=1}^{m^0+1}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=T_{k-1}^0}^{T_k^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}X_s\tilde{\eta}_k f_s^{0\prime}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^{0}\right)^{+} f_t^0 \\
= \;& -\sum_{k=1}^{m^0+1}\left(\frac{1}{NT\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\chi_{st}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}X_s\right)\tilde{\eta}_k \\
= \;& -\left[\tilde{\Phi}_{j1}^*(\tilde{\boldsymbol{\Lambda}}),\dots,\tilde{\Phi}_{j,m^0+1}^*(\tilde{\boldsymbol{\Lambda}})\right]\left(\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\right), \qquad \text{(C.98)}
\end{aligned}$$

where $\chi_{st} = f_s^{0\prime}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^{0}\right)^{+}f_t^0$ and $\tilde{\Phi}_{jk}^*(\tilde{\boldsymbol{\Lambda}}) = \frac{1}{NT\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\chi_{st}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}X_s$. By Lemmas C.4(v) and C.5(i), we may show that

$$\left\|\tilde{\Phi}_{jk}^*(\tilde{\boldsymbol{\Lambda}}) - \Phi_{jk}^*\right\| = O_P\left(p\delta_{p,NT}^{-1}(m^0)^{-1}\right), \quad 1\le j,k\le m^0+1, \qquad \text{(C.99)}$$

where $\Phi^*_{jk}=\frac{1}{NT\tau_j(T)}\sum_{t=T^0_{j-1}}^{T^0_j-1}\sum_{s=T^0_{k-1}}^{T^0_k-1}\chi_{st}X_t'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}X_s$. Hence, by (C.98), (C.99) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\left\|\frac{1}{N\tau_j(T)}\sum_{t=T^0_{j-1}}^{T^0_j-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,2}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}_{m^0}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+f^0_t+\left(\Phi^*_{j1},...,\Phi^*_{j,m^0+1}\right)(\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0)\right\| \\
= & \left\|\left[\tilde{\Phi}^*_{j1}(\tilde{\boldsymbol{\Lambda}}),...,\tilde{\Phi}^*_{j,m^0+1}(\tilde{\boldsymbol{\Lambda}})\right](\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0)-\left(\Phi^*_{j1},...,\Phi^*_{j,m^0+1}\right)(\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0)\right\| \\
= & O_P\left(p\delta^{-1}_{p,NT}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right). \qquad \text{(C.100)}
\end{aligned}$$

For $l=3$, by the definition of $\tilde{u}_{NT,3}$, Assumptions 1 and 3(ii), as well as (C.92), we have

$$\begin{aligned}
&\frac{1}{N\tau_j(T)}\left\|\sum_{t=T^0_{j-1}}^{T^0_j-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,3}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+f^0_t\right\| \\
= & \frac{1}{N\tau_j(T)}\left\|\sum_{t=T^0_{j-1}}^{T^0_j-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\left(\frac{1}{NT}\sum_{k=1}^{m^0+1}\sum_{s=T^0_{k-1}}^{T^0_k-1}X_s\tilde{\eta}_k\varepsilon_s'\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+f^0_t\right\| \\
= & \frac{O_P(1)}{N^2T\tau_j(T)}\sum_{k=1}^{m^0+1}\|\tilde{\eta}_k\|\left[\sum_{t=T^0_{j-1}}^{T^0_j-1}\sum_{s=T^0_{k-1}}^{T^0_k-1}\left\|X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}X_s\right\|\left(\left\|\varepsilon_s'\boldsymbol{\Lambda}^0\right\|+\left\|\varepsilon_s'(\tilde{\boldsymbol{\Lambda}}-\boldsymbol{\Lambda}^0\tilde{\boldsymbol{H}})\right\|\right)\left\|f^0_t\right\|\right] \\
= & O_P\left(p\delta^{-1}_{NT}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right). \qquad \text{(C.101)}
\end{aligned}$$

To study the next two terms, we can apply the arguments used in the proof of Lemma C.3(ii) and show that $\frac{1}{N}||X_t'(\boldsymbol{\Lambda}^0-\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}^+)||=O_P(p^{1/2}\delta^{-2}_{NT}+\tilde{\eta}^{1/2}_{NT})$. This, in conjunction with Lemma C.4(iii), implies that

$$\frac{1}{N}\left\|X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}^+)\right\|=O_P\left(p^{1/2}\delta^{-2}_{NT}+\tilde{\eta}^{1/2}_{NT}\right) \qquad \text{(C.102)}$$

and similarly for $j=1,\cdots,m^0+1$,

$$\frac{1}{N\tau_j(T)}\sum_{t=T^0_{j-1}}^{T^0_j-1}\left\|X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}(\boldsymbol{\Lambda}^0-\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{H}}^+)\right\|\left\|f^0_t\right\|=O_P\left(p^{1/2}\delta^{-2}_{NT}+\tilde{\eta}^{1/2}_{NT}\right). \qquad \text{(C.103)}$$

For $l = 4$, by the definition of $\tilde{u}_{NT,4}$, (C.103), and Lemma C.5(i) and noting that $\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{\boldsymbol{\Lambda}} = \boldsymbol{0}$,

$$\begin{aligned}
& \frac{1}{N\tau_j(T)} \left\| \sum_{t=T_{j-1}^0}^{T_j^0-1} X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \tilde{u}_{NT,4} \left(\frac{1}{N} \boldsymbol{\Lambda}^{0\prime} \tilde{\boldsymbol{\Lambda}} \right)^+ \left(\frac{1}{T} \boldsymbol{F}^{0\prime} \boldsymbol{F}^0 \right)^+ f_t^0 \right\| \\
= \; & \frac{1}{N\tau_j(T)} \left\| \sum_{t=T_{j-1}^0}^{T_j^0-1} X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 - \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{H}}^+) \left(\frac{1}{NT} \sum_{k=1}^{m^0+1} \sum_{s=T_{k-1}^0}^{T_k^0-1} f_s^0 \tilde{\eta}_k' X_s' \tilde{\boldsymbol{\Lambda}} \right) \left(\frac{1}{N} \boldsymbol{\Lambda}^{0\prime} \tilde{\boldsymbol{\Lambda}} \right)^+ \left(\frac{1}{T} \boldsymbol{F}^{0\prime} \boldsymbol{F}^0 \right)^+ f_t^0 \right\| \\
= \; & O_P \left(\frac{1}{N^2 T \tau_j(T)} \sum_{k=1}^{m^0+1} \|\tilde{\eta}_k\| \sum_{t=T_{j-1}^0}^{T_j^0-1} \sum_{s=T_{k-1}^0}^{T_k^0-1} \|X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 - \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{H}}^+)\| \|X_s' \tilde{\boldsymbol{\Lambda}}\| \|f_s^0\| \|f_t^0\| \right) \\
= \; & O_P \left(\delta_{p,NT}^{-1} (m^0)^{-1/2} \|\tilde{\boldsymbol{\alpha}}_{m^0} - \boldsymbol{\alpha}^0\| \right). \qquad \text{(C.104)}
\end{aligned}$$

For $l = 5$, by the definition of $\tilde{u}_{NT,5}$, Assumptions 1(i)(iii), (C.103), and Lemma C.5(iv), we have

$$\begin{aligned}
& \frac{1}{N\tau_j(T)} \left\| \sum_{t=T_{j-1}^0}^{T_j^0-1} X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \tilde{u}_{NT,5} \left(\frac{1}{N} \boldsymbol{\Lambda}^{0\prime} \tilde{\boldsymbol{\Lambda}} \right)^+ \left(\frac{1}{T} \boldsymbol{F}^{0\prime} \boldsymbol{F}^0 \right)^+ f_t^0 \right\| \\
= \; & \frac{1}{N\tau_j(T)} \left\| \sum_{t=T_{j-1}^0}^{T_j^0-1} X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \left(\frac{1}{NT} \sum_{s=1}^{T} \boldsymbol{\Lambda}^0 f_s^0 \varepsilon_s' \tilde{\boldsymbol{\Lambda}} \right) \left(\frac{1}{N} \boldsymbol{\Lambda}^{0\prime} \tilde{\boldsymbol{\Lambda}} \right)^+ \left(\frac{1}{T} \boldsymbol{F}^{0\prime} \boldsymbol{F}^0 \right)^+ f_t^0 \right\| \\
\leq \; & \frac{1}{N^2 T \tau_j(T)} \left\| \sum_{t=T_{j-1}^0}^{T_j^0-1} X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 - \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{H}}^+) \left(\sum_{s=1}^{T} f_s^0 \varepsilon_s' \boldsymbol{\Lambda}^0 \right) \tilde{\boldsymbol{H}} \left(\frac{1}{N} \boldsymbol{\Lambda}^{0\prime} \tilde{\boldsymbol{\Lambda}} \right)^+ \left(\frac{1}{T} \boldsymbol{F}^{0\prime} \boldsymbol{F}^0 \right)^+ f_t^0 \right\| \\
& + \frac{1}{N^2 T \tau_j(T)} \left\| \sum_{t=T_{j-1}^0}^{T_j^0-1} X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 - \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{H}}^+) \left[\sum_{s=1}^{T} f_s^0 \varepsilon_s' (\tilde{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0 \tilde{\boldsymbol{H}}) \right] \left(\frac{1}{N} \boldsymbol{\Lambda}^{0\prime} \tilde{\boldsymbol{\Lambda}} \right)^+ \left(\frac{1}{T} \boldsymbol{F}^{0\prime} \boldsymbol{F}^0 \right)^+ f_t^0 \right\| \\
= \; & O_P \left(\frac{1}{N^2 T \tau_j(T)} \sum_{t=T_{j-1}^0}^{T_j^0-1} \|X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 - \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{H}}^+)\| \left\| \sum_{s=1}^{T} f_s^0 \varepsilon_s' \boldsymbol{\Lambda}^0 \right\| \|f_t^0\| \right) \\
& + O_P \left(\frac{1}{N^2 T \tau_j(T)} \sum_{t=T_{j-1}^0}^{T_j^0-1} \sum_{s=1}^{T} \|X_t' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} (\boldsymbol{\Lambda}^0 - \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{H}}^+)\| \|\varepsilon_s' (\tilde{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}^0 \tilde{\boldsymbol{H}}) f_s^0\| \|f_t^0\| \right) \\
= \; & O_P \left(\delta_{p,NT}^{-3} + \delta_{p,NT}^{-2} (m^0)^{-1/2} \|\tilde{\boldsymbol{\alpha}}_{m^0} - \boldsymbol{\alpha}^0\| \right). \qquad \text{(C.105)}
\end{aligned}$$

For $l=6$, by the definition of $\tilde{u}_{NT,6}$ and Assumptions 1(i)-(iii), 2(ii) and 3(ii), we have

$$
\begin{aligned}
& \frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,6}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+ f_t^0\right\| \\
= & \frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\left(\frac{1}{NT}\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\varepsilon_s\tilde{\eta}_k' X_s'\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+ f_t^0\right\| \\
\leq & \frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1}\left(\frac{1}{NT}\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1} X_t'\varepsilon_s\tilde{\eta}_k' X_s'\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+ f_t^0\right\| \\
& +\frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\left(\frac{1}{NT}\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}(\boldsymbol{P}_{\tilde{\boldsymbol{\Lambda}}}-\boldsymbol{P}_{\boldsymbol{\Lambda}^0})\varepsilon_s\tilde{\eta}_k' X_s'\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+ f_t^0\right\| \\
& +\frac{1}{N\tau_j(T)}\left\|\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{\Lambda}^0(\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0)^+\left(\frac{1}{NT}\sum_{k=1}^{m^0+1}\sum_{s=T_{k-1}^0}^{T_k^0-1}\boldsymbol{\Lambda}^{0\prime}\varepsilon_s\tilde{\eta}_k' X_s'\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+ f_t^0\right\| \\
= & O_P\left(p^{1/2}\delta_{p,NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m^0}-\boldsymbol{\alpha}^0\|\right). \qquad \text{(C.106)}
\end{aligned}
$$

For $l=7$, by the definitions of $\tilde{u}_{NT,7}$ and χ_{st}, we have

$$
\begin{aligned}
& \frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,7}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+ f_t^0 \\
= & \frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\left(\frac{1}{NT}\sum_{s=1}^{T}\varepsilon_s f_s^{0\prime}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^+\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^+ f_t^0 \\
= & \frac{1}{NT\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=1}^{T}\chi_{st}X_t'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\varepsilon_s+\frac{1}{NT\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=1}^{T}\chi_{st}X_t'(\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}-\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\varepsilon_s \\
= & \frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1} X_t'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\varepsilon_t^*+\frac{1}{NT\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=1}^{T}\chi_{st}X_t'(\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}-\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\varepsilon_s, \qquad \text{(C.107)}
\end{aligned}
$$

where $\varepsilon_t^*=\frac{1}{T}\sum_{s=1}^T\chi_{st}\varepsilon_s$. On the other hand, following the proof of Lemma B.3(i) and (C.95) in particular, we may show that $\left\|\frac{1}{NT\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}\sum_{s=1}^{T}\chi_{st}X_t'(\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}-\boldsymbol{M}_{\boldsymbol{\Lambda}^0})\varepsilon_s+B_{NT,j}(2,2)\right\| =$

$O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m_0}-\boldsymbol{\alpha}^0\|+\delta_{p,NT}^{-3}\right)$. It follows that

$$\begin{aligned}
&\left\|\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,7}\big(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\big)^{+}\big(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\big)^{+}f_t^0-\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\boldsymbol{M}_{\boldsymbol{\Lambda}^0}\varepsilon_t^*+B_{NT,j}(2,2)\right\| \\
=\;& O_P\left(\delta_{NT}^{-1}(m^0)^{-1/2}\|\tilde{\boldsymbol{\alpha}}_{m_0}-\boldsymbol{\alpha}^0\|+\delta_{p,NT}^{-3}\right). \qquad (C.108)
\end{aligned}$$

For $l=8$, by the definition of $\tilde{u}_{NT,8}$, we have

$$\begin{aligned}
&\frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\tilde{u}_{NT,8}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{+}f_t^0 \\
=\;& \frac{1}{N\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\left(\frac{1}{NT}\sum_{s=1}^{T}\varepsilon_s\varepsilon_s'\tilde{\boldsymbol{\Lambda}}\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{+}f_t^0 \\
=\;& \frac{1}{N^2T\tau_j(T)}\sum_{t=T_{j-1}^0}^{T_j^0-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\tilde{\boldsymbol{\Lambda}}\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\tilde{\boldsymbol{\Lambda}}\right)^{+}\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)^{+}f_t^0\equiv B_{NT,j}(1). \qquad (C.109)
\end{aligned}$$

By (C.96), (C.97), (C.100), (C.101), (C.104)–(C.106), (C.108) and (C.109), we can complete the proof of Lemma B.3(ii).

We have thus completed the proof of Lemma B.3. ■

Let $\dot{\boldsymbol{\Lambda}}_R=(\dot{\lambda}_{1,R},...,\dot{\lambda}_{N,R})'$ and $\breve{\boldsymbol{\Lambda}}_R=\frac{1}{NT}\sum_{t=1}^{T}(Y_t-X_t\dot{\beta}_{t,R})(Y_t-X_t\dot{\beta}_{t,R})'\dot{\boldsymbol{\Lambda}}_R=(\breve{\lambda}_{1,R},...,\breve{\lambda}_{N,R})'$. In order to prove Lemma B.4 in Appendix B, we first need to prove the following technical lemma.

Lemma C.6 *Suppose that Assumptions 1 and 2 in Appendix A hold and $R>R_0$. Define the $R_0\times R$ matrix $\dot{\boldsymbol{H}}_R\equiv\left(\frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\right)\left(\frac{1}{N}\boldsymbol{\Lambda}^{0\prime}\dot{\boldsymbol{\Lambda}}_R\right)$ with the Moore-Penrose generalized inverse $\dot{\boldsymbol{H}}_R^{+}=\begin{bmatrix}\dot{\boldsymbol{H}}_R^{+}(1)\\ \dot{\boldsymbol{H}}_R^{+}(2)\end{bmatrix}$, where $\dot{\boldsymbol{H}}_R^{+}(1)$ and $\dot{\boldsymbol{H}}_R^{+}(2)$ are $R_0\times R_0$ and $(R-R_0)\times R_0$ matrices, respectively. Let $\dot{\boldsymbol{V}}_{NT,R}$ denote an $R\times R$ diagonal matrix consisting of the R largest eigenvalues of the $N\times N$ matrix $\frac{1}{NT}\sum_{t=1}^{T}(Y_t-X_t\dot{\beta}_{t,R})(Y_t-X_t\dot{\beta}_{t,R})'$ where the eigenvalues are in decreasing order along the main diagonal line. Write $\dot{\boldsymbol{\Lambda}}_R=\left[\dot{\boldsymbol{\Lambda}}_R(1),\dot{\boldsymbol{\Lambda}}_R(2)\right]$ and $\dot{\boldsymbol{H}}_R=\left[\dot{\boldsymbol{H}}_R(1),\dot{\boldsymbol{H}}_R(2)\right]$, where $\dot{\boldsymbol{\Lambda}}_R(1)$, $\dot{\boldsymbol{\Lambda}}_R(2)$, $\dot{\boldsymbol{H}}_R(1)$, and $\dot{\boldsymbol{H}}_R(2)$ are $N\times R_0$, $N\times(R-R_0)$, $R_0\times R_0$, and $R_0\times(R-R_0)$ matrices, respectively. Furthermore, write $\dot{\boldsymbol{V}}_{NT,R}=\mathsf{diag}\left\{\dot{\boldsymbol{V}}_{NT,R}(1),\dot{\boldsymbol{V}}_{NT,R}(2)\right\}$, where $\dot{\boldsymbol{V}}_{NT,R}(1)$ denotes the upper-left $R_0\times R_0$ submatrix of $\dot{\boldsymbol{V}}_{NT,R}$. Then we have*

(i) $\frac{1}{N}\left\|\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R\right\|^2=O_P\left(\delta_{p,NT}^{-2}\right)$,

(ii) $\frac{1}{N}\left\|\breve{\boldsymbol{\Lambda}}_R'\breve{\boldsymbol{\Lambda}}_R-\dot{\boldsymbol{H}}_R'\boldsymbol{\Lambda}^{0\prime}\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R\right\|=O_P\left(\delta_{p,NT}^{-1}\right)$,

(iii) $\frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R(1)-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(1)\dot{\boldsymbol{V}}_{NT,R}^{+}(1)\right\|^2=O_P\left(\delta_{p,NT}^{-2}\right)$ *and* $\left\|\dot{\boldsymbol{H}}_R(2)\right\|^2=O_P\left(\delta_{p,NT}^{-2}\right)$,

(iv) $\left\|\dot{\boldsymbol{H}}_R^{+}(1)\right\|=O_P(1)$ *and* $\left\|\dot{\boldsymbol{H}}_R^{+}(2)\right\|=O_P\left(\delta_{p,NT}^{-1}\right)$.

Proof of Lemma C.6. **(i)** When $R>R_0$, we can follow the proof of Lemma C.2 and show that $\dot{\eta}_R\equiv\frac{1}{T}\sum_{t=1}^T\|\dot{\beta}_{t,R}-\beta_t^0\|^2=o_P(1)$. Next, using $Y_t-X_t\dot{\beta}_{t,R}=\mathbf{\Lambda}^0f_t^0+\varepsilon_t+X_t(\beta_t^0-\dot{\beta}_{t,R})$ and $\dot{d}_{t,R}=\dot{\beta}_{t,R}-\beta_t^0$, we have

$$\begin{aligned}
\breve{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R &= \frac{1}{NT}\sum_{t=1}^T(Y_t-X_t\dot{\beta}_{t,R})(Y_t-X_t\dot{\beta}_{t,R})'\dot{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R\\
&= \frac{1}{NT}\sum_{t=1}^T\left[-X_t\dot{d}_{t,R}+\mathbf{\Lambda}^0f_t^0+\varepsilon_t\right]\left[-X_t\dot{d}_{t,R}+\mathbf{\Lambda}^0f_t^0+\varepsilon_t\right]'\dot{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R\\
&= \frac{1}{NT}\sum_{t=1}^T X_t\dot{d}_{t,R}\dot{d}_{t,R}'X_t'\dot{\mathbf{\Lambda}}_R-\frac{1}{NT}\sum_{t=1}^T X_td_{t,R}f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R-\frac{1}{NT}\sum_{t=1}^T X_t\dot{d}_{t,R}\varepsilon_t'\dot{\mathbf{\Lambda}}_R\\
&\quad -\frac{1}{NT}\sum_{t=1}^T\mathbf{\Lambda}^0f_t^0\dot{d}_{t,R}'X_t'\dot{\mathbf{\Lambda}}_R+\frac{1}{NT}\sum_{t=1}^T\mathbf{\Lambda}^0f_t^0\varepsilon_t'\dot{\mathbf{\Lambda}}_R-\frac{1}{NT}\sum_{t=1}^T\varepsilon_t\dot{d}_{t,R}'X_t'\dot{\mathbf{\Lambda}}_R\\
&\quad +\frac{1}{NT}\sum_{t=1}^T\varepsilon_tf_t^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R+\frac{1}{NT}\sum_{t=1}^T\varepsilon_t\varepsilon_t'\dot{\mathbf{\Lambda}}_R\\
&\equiv \sum_{j=1}^8\dot{u}_{R,j}.
\end{aligned}\tag{C.110}$$

Following the proof of Lemma C.3(i), we can readily show that $\frac{1}{N}||\dot{u}_{R,j}||^2=O_P\left(\delta_{NT}^{-2}+\dot{\eta}_R\right).$ Then we readily have $\frac{1}{N}||\breve{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R||^2=O_P\left(\delta_{NT}^{-2}+\dot{\eta}_R\right)$. With this, we can apply the arguments used in the proof of Theorem 3.1 to show that $\dot{\eta}_R=O_P\left(\delta_{p,NT}^{-2}\right)$. Then we may complete the proof of (i).

(ii) Noting that

$$\begin{aligned}
&\frac{1}{N}\breve{\mathbf{\Lambda}}_R'\breve{\mathbf{\Lambda}}_R-\frac{1}{N}\dot{\boldsymbol{H}}_R'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R\\
=\ &\frac{1}{N}(\breve{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)'(\breve{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)+\frac{1}{N}(\breve{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)'\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R+\frac{1}{N}\dot{\boldsymbol{H}}_R'\mathbf{\Lambda}^{0\prime}(\breve{\mathbf{\Lambda}}_R-\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R),
\end{aligned}$$

the convergence result (ii) follows from the triangle and Cauchy-Schwarz inequalities, Lemma C.6(i), and the fact that $||\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R||^2=O_P(N)$.

(iii) Let $\dot{\boldsymbol{V}}_R$ and $\dot{\boldsymbol{V}}_R(1)$ denote the probability limits of $\dot{\boldsymbol{V}}_{NT,R}$ and $\dot{\boldsymbol{V}}_{NT,R}(1)$, respectively, as $(N,T)\to\infty$. Recall that $\dot{\boldsymbol{H}}_R=\frac{1}{NT}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R$ and $\frac{1}{N}\dot{\mathbf{\Lambda}}_R'\dot{\mathbf{\Lambda}}_R=\boldsymbol{I}_R$. As the application of

PCA method, we have the identity

$$\frac{1}{NT}\sum_{t=1}^{T}(Y_t - X_t\dot{\beta}_{t,R})(Y_t - X_t\dot{\beta}_{t,R})'\dot{\mathbf{\Lambda}}_R = \dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R}.$$

Pre-multiplying both sides of the above equation by $\dot{\mathbf{\Lambda}}_R'/N$ and using the normalization $\frac{1}{N}\dot{\mathbf{\Lambda}}_R'\dot{\mathbf{\Lambda}}_R = \boldsymbol{I}_R$ yields

$$\frac{1}{N^2T}\dot{\mathbf{\Lambda}}_R'\left[\sum_{t=1}^{T}(Y_t - X_t\dot{\beta}_{t,R})(Y_t - X_t\dot{\beta}_{t,R})'\right]\dot{\mathbf{\Lambda}}_R = \dot{\mathbf{V}}_{NT,R},$$

which together with $Y_t - X_t\dot{\beta}_{t,R} = X_t(\beta_t^0 - \dot{\beta}_{t,R}) + \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t$, yields $\frac{1}{N^2T}\dot{\mathbf{\Lambda}}_R'\mathbf{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R + d_{NT,R} = \dot{\mathbf{V}}_{NT,R}$, where

$$\begin{aligned}
d_{NT,R} &= \frac{1}{N^2T}\dot{\mathbf{\Lambda}}_R'\sum_{t=1}^{T}\Big[X_t(\beta_t^0 - \dot{\beta}_{t,R})(\beta_t^0 - \dot{\beta}_{t,R})'X_t' + \varepsilon_t\varepsilon_t' + X_t(\beta_t^0 - \dot{\beta}_{t,R})f_t^{0\prime}\mathbf{\Lambda}^{0\prime} \\
&\quad + \mathbf{\Lambda}^0 f_t^0(\beta_t^0 - \dot{\beta}_{t,R})'X_t' + X_t(\beta_t^0 - \dot{\beta}_{t,R})\varepsilon_t' + \varepsilon_t(\beta_t^0 - \dot{\beta}_{t,R})'X_t' \\
&\quad + \mathbf{\Lambda}^0 f_t^0\varepsilon_t' + \varepsilon_t f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\Big]\dot{\mathbf{\Lambda}}_R \\
&\equiv \sum_{j=1}^{8} d_{R,j}.
\end{aligned}$$

Following the proof of Lemma C.3, it is easy to show that $\|d_{NT,R}\| = O_P\left(\delta_{p,NT}^{-1}\right)$ by proving that $d_{R,j}$, $j = 1, 2, ..., 8$, are either $O_P(\delta_{p,NT}^{-1})$ or of smaller order. For example,

$$\begin{aligned}
\|d_{R,1}\| &= \frac{1}{N^2T}\left\|\dot{\mathbf{\Lambda}}_R'\left[\sum_{t=1}^{T}X_t(\beta_t^0 - \dot{\beta}_{t,R})(\beta_t^0 - \dot{\beta}_{t,R})'X_t'\right]\dot{\mathbf{\Lambda}}_R\right\| \\
&\le \frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\right\|^2 \mu_{\max}\left(X_t'X_t/N\right)\frac{1}{T}\sum_{t=1}^{T}\left\|\beta_t^0 - \dot{\beta}_{t,R}\right\|^2 = O_P\left(\delta_{p,NT}^{-2}\right), \\
\|d_{R,2}\| &= \frac{1}{N^2T}\left\|\dot{\mathbf{\Lambda}}_R'\left[\sum_{t=1}^{T}\varepsilon_t\varepsilon_t'\right]\dot{\mathbf{\Lambda}}_R\right\| \le \frac{1}{NT}\|\boldsymbol{\varepsilon}\|_{\text{sp}}^2\frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\right\|^2 = O_P\left(\delta_{NT}^{-2}\right),
\end{aligned}$$

and

$$\begin{aligned}
\|d_{R,3}\| &= \frac{1}{N^2T}\left\|\dot{\mathbf{\Lambda}}_R'\sum_{t=1}^{T}X_t(\beta_t^0 - \dot{\beta}_{t,R})f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R\right\| \\
&\le \frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\right\|^2\frac{1}{N^{1/2}}\left\|\mathbf{\Lambda}^0\right\|\mu_{\max}^{1/2}\left(X_t'X_t/N\right)\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\beta_t^0 - \dot{\beta}_{t,R}\right\|^2\right)^{1/2}\frac{1}{T^{1/2}}\left\|\mathbf{F}^0\right\| \\
&\le O_P\left(\dot{\eta}_R^{1/2}\right) = O_P\left(\delta_{p,NT}^{-1}\right).
\end{aligned}$$

Then

$$\frac{1}{N^2T}\dot{\mathbf{\Lambda}}_R'\mathbf{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R = \dot{\mathbf{V}}_{NT,R} - d_{NT,R} \xrightarrow{P} \dot{\mathbf{V}}_R. \tag{C.111}$$

Observe that $\frac{1}{N^2T}\dot{\mathbf{\Lambda}}_R'\mathbf{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R$ has rank R_0 at most in both finite and large samples. Let $\mathbf{\Delta}_{NT}(l) = \frac{1}{N}\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R(l)$ for $l = 1, 2$ and $\hat{\mathbf{\Sigma}}_F = \frac{1}{T}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0$. Then

$$\frac{1}{N^2T}\dot{\mathbf{\Lambda}}_R'\mathbf{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R = \begin{bmatrix} \mathbf{\Delta}_{NT}'(1)\hat{\mathbf{\Sigma}}_F\mathbf{\Delta}_{NT}(1) & \mathbf{\Delta}_{NT}'(1)\hat{\mathbf{\Sigma}}_F\mathbf{\Delta}_{NT}(2) \\ \mathbf{\Delta}_{NT}'(2)\hat{\mathbf{\Sigma}}_F\mathbf{\Delta}_{NT}(1) & \mathbf{\Delta}_{NT}'(2)\hat{\mathbf{\Sigma}}_F\mathbf{\Delta}_{NT}(2) \end{bmatrix}.$$

Note that $\hat{\mathbf{\Sigma}}_F = \mathbf{\Sigma}_F + o_P(1)$ by Assumption 1(i). Following the proof of Lemma A.3(ii) in Bai (2003), we can show that $\text{plim}_{(N,T)\to\infty}\mathbf{\Delta}_{NT}'(1)\hat{\mathbf{\Sigma}}_F\mathbf{\Delta}_{NT}(1) = \dot{\mathbf{V}}_R(1)$ which has full rank R_0. This ensures that $\frac{1}{N^2T}\dot{\mathbf{\Lambda}}_R'\mathbf{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R$ has rank R_0 in large samples and $\mathbf{\Delta}_{NT}'(2)\hat{\mathbf{\Sigma}}_F\mathbf{\Delta}_{NT}(2) \xrightarrow{P} \mathbf{0}$. Then $\mathbf{\Delta}_{NT}'(1)\hat{\mathbf{\Sigma}}_F\mathbf{\Delta}_{NT}(2) \xrightarrow{P} \mathbf{0}$ by the Cauchy-Schwarz inequality. By the asymptotic nonsingularity of $\hat{\mathbf{\Sigma}}_F$, this also implies that $\mathbf{\Delta}_{NT}(2) = o_P(1)$ and $\mathbf{\Delta}_{NT}(1) \xrightarrow{P} \mathbf{\Delta}(1)$ for some $R_0 \times R_0$ nonsingular matrix $\mathbf{\Delta}(1)$. Consequently, we have

$$\dot{\boldsymbol{H}}_R(1) = \frac{1}{NT}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R(1) \xrightarrow{P} \mathbf{\Sigma}_F\mathbf{\Delta}(1)$$

and

$$\dot{\boldsymbol{H}}_R(2) = \frac{1}{NT}\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R(2) = o_P(1).$$

Then $\dot{\boldsymbol{H}}_R(1)$ is asymptotically nonsingular and $\dot{\boldsymbol{H}}_R$ has rank R_0.

By the definition $\breve{\mathbf{\Lambda}}_R = \frac{1}{NT}\sum_{t=1}^T(Y_t - X_t\dot{\beta}_{t,R})(Y_t - X_t\dot{\beta}_{t,R})'\dot{\mathbf{\Lambda}}_R$ and the identity $\frac{1}{NT}\sum_{t=1}^T (Y_t - X_t\dot{\beta}_{t,R})(Y_t - X_t\dot{\beta}_{t,R})'\dot{\mathbf{\Lambda}}_R = \dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R}$ from the PCA, we have

$$\begin{aligned}\frac{1}{N}\left\|\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R\right\|^2 &= \frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R} - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R\right\|^2 \\ &= \frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R}(1) - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(1)\right\|^2 + \frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R}(2) - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(2)\right\|^2.\end{aligned}$$

Lemma C.6(i) implies that $\frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R}(l) - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(l)\right\|^2 = O_P(\delta_{p,NT}^{-2})$ for $l = 1, 2$. Since $\dot{\mathbf{V}}_R(1)$ is nonsingular, it follows that $\frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(1)\dot{\mathbf{V}}_{NT,R}^+(1)\right\|^2 = O_P(\delta_{p,NT}^{-2})$ and $\left\|\dot{\mathbf{V}}_{NT,R}^+(1)\right\| \leq \left\|\dot{\mathbf{V}}_R^+(1)\right\| + \left\|\dot{\mathbf{V}}_{NT,R}^+(1) - \dot{\mathbf{V}}_R^+(1)\right\| = O_P(1)$.

In addition,

$$\begin{aligned}\frac{1}{N}\left\|\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(2)\right\|^2 &\leq \frac{2}{N}\left\|\dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R}(2) - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(2)\right\|^2 + \frac{2}{N}\left\|\dot{\mathbf{\Lambda}}_R\dot{\mathbf{V}}_{NT,R}(2)\right\|^2 \\ &= O_P\left(\delta_{p,NT}^{-2}\right) + O_P\left(\delta_{p,NT}^{-2}\right) = O_P\left(\delta_{p,NT}^{-2}\right),\end{aligned}$$

because $\frac{1}{N}\left\|\dot{\mathbf{\Lambda}}_R\dot{\boldsymbol{V}}_{NT,R}(2)\right\|^2 \leq \mu^2_{\max}(\dot{\boldsymbol{V}}_{NT,R}(2))\left\|\dot{\mathbf{\Lambda}}_R\right\|^2/N = R\mu^2_{\max}(\dot{\boldsymbol{V}}_{NT,R}(2))$ and $\mu_{\max}(\dot{\boldsymbol{V}}_{NT,R}(2)) \leq \mu_{R_0+1}(\dot{\mathbf{\Lambda}}_R'\mathbf{\Lambda}^0\boldsymbol{F}^{0\prime}\boldsymbol{F}^0\mathbf{\Lambda}^{0\prime}\dot{\mathbf{\Lambda}}_R/(N^2T)) + \|d_{NT,R}\| = \|d_{NT,R}\| = O_P(\delta^{-1}_{p,NT})$, where $\mu_{R_0+1}(\cdot)$ denotes the (R_0+1)-th largest eigenvalue of the square matrix in the parentheses. In view of the fact that

$$\frac{1}{N}\left\|\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(2)\right\|^2 = \frac{1}{N}\mathsf{Tr}\left(\dot{\boldsymbol{H}}_R(2)\dot{\boldsymbol{H}}_R(2)'\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0\right) \geq \mu_{\min}\left(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0/N\right)\left\|\dot{\boldsymbol{H}}_R(2)\right\|^2,$$

we have $\left\|\dot{\boldsymbol{H}}_R(2)\right\|^2 \leq \left[\mu_{\min}\left(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0/N\right)\right]^{-1}\frac{1}{N}\left\|\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R(2)\right\|^2 = O_P(\delta^{-2}_{p,NT})$.

(iv) Since $\dot{\boldsymbol{H}}_R$ is right invertible asymptotically, by Proposition 6.1.5 in Bernstein (2005, p.225), the $R\times R_0$ generalized inverse $\dot{\boldsymbol{H}}_R^+$ of $\dot{\boldsymbol{H}}_R$ is given by

$$\dot{\boldsymbol{H}}_R^+ = \dot{\boldsymbol{H}}_R'\left[\dot{\boldsymbol{H}}_R\dot{\boldsymbol{H}}_R'\right]^{-1} = \begin{bmatrix} \dot{\boldsymbol{H}}_R'(1)\left(\dot{\boldsymbol{H}}_R\dot{\boldsymbol{H}}_R'\right)^{-1} \\ \dot{\boldsymbol{H}}_R'(2)\left(\dot{\boldsymbol{H}}_R\dot{\boldsymbol{H}}_R'\right)^{-1}\end{bmatrix} = \begin{bmatrix}\dot{\boldsymbol{H}}_R^+(1) \\ \dot{\boldsymbol{H}}_R^+(2)\end{bmatrix}.$$

Then by Lemma C.6(iii)

$$\begin{aligned}\left\|\dot{\boldsymbol{H}}_R^+(1)\right\| &\leq \left\|\dot{\boldsymbol{H}}_R(1)\right\|\left\|\left(\dot{\boldsymbol{H}}_R\dot{\boldsymbol{H}}_R'\right)^{-1}\right\| = O_P(1), \text{ and} \\ \left\|\dot{\boldsymbol{H}}_R^+(2)\right\| &\leq \left\|\dot{\boldsymbol{H}}_R(2)\right\|\left\|\left(\dot{\boldsymbol{H}}_R\dot{\boldsymbol{H}}_R'\right)^{-1}\right\| = O_P\left(\delta^{-1}_{p,NT}\right).\end{aligned}$$

We have thus completed the proof of Lemma C.6. ■

Proof of Lemma B.4. **(i)** The proof is similar to that of Lemma C.2. Notice that

$$\hat{Q}_{NT}(\boldsymbol{\beta},\mathbf{\Lambda}_R) = \frac{1}{NT}\sum_{t=1}^T(Y_t - X_t\beta_t)'\boldsymbol{M}_{\mathbf{\Lambda}_R}(Y_t - X_t\beta_t).$$

Using $Y_t - X_t\dot{\beta}_{t,R} = X_t(\beta^0_t - \dot{\beta}_{t,R}) + \mathbf{\Lambda}^0 f^0_t + \varepsilon_t$, we have

$$\begin{aligned} 0 &\geq \hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_R,\dot{\mathbf{\Lambda}}_R) - \hat{Q}_{NT}(\boldsymbol{\beta}^0,\dot{\mathbf{\Lambda}}_R) \\ &= \frac{1}{NT}\sum_{t=1}^T\left[(Y_t - X_t\dot{\beta}_{t,R})'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}(Y_t - X_t\dot{\beta}_{t,R}) - (Y_t - X_t\beta^0_t)'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}(Y_t - X_t\beta^0_t)\right] \\ &= \frac{1}{NT}\sum_{t=1}^T\left[(\dot{\beta}_{t,R} - \beta^0_t)'X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_t(\dot{\beta}_{t,R} - \beta^0_t) - 2(\dot{\beta}_{t,R} - \beta^0_t)'X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\mathbf{\Lambda}^0 f^0_t\right] \\ &\quad -\frac{2}{NT}\sum_{t=1}^T(\dot{\beta}_{t,R} - \beta^0_t)'X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\varepsilon_t. \end{aligned}$$

By Lemma C.1(i) (with R_0 and $\mathbf{\Lambda}$ being replaced by R and $\mathbf{\Lambda}_R$), we can prove that

$$\frac{1}{NT}\sum_{t=1}^{T}(\dot{\beta}_{t,R}-\beta_t^0)'X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\varepsilon_t = O_P\left(p^{1/2}\delta_{p,NT}^{-1}\right).$$

Let $\dot{\boldsymbol{d}}_{\beta,R} = \dot{\boldsymbol{\beta}}_R - \boldsymbol{\beta}^0$ and $\dot{\boldsymbol{d}}_{\Lambda,R} = \frac{1}{N^{1/2}}\mathsf{vec}(\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\mathbf{\Lambda}^0)$. Define

$$\dot{\boldsymbol{A}}_R = \frac{1}{N}\mathsf{diag}(X_1'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_1, ..., X_T'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_T) \text{ and } \dot{\boldsymbol{C}}_R = \frac{1}{N^{1/2}}[f_1^0\otimes\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_1, ..., f_T^0\otimes\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_T].$$

Then

$$\begin{aligned}&\frac{1}{NT}\sum_{t=1}^{T}\left[(\dot{\beta}_{t,R}-\beta_t^0)'X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_t(\dot{\beta}_{t,R}-\beta_t^0) - 2(\dot{\beta}_{t,R}-\beta_t^0)'X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\mathbf{\Lambda}^0 f_t^0\right]\\ =\;& \frac{1}{T}\dot{\boldsymbol{d}}_{\beta,R}'\dot{\boldsymbol{A}}_R\dot{\boldsymbol{d}}_{\beta,R} - \frac{2}{T}\dot{\boldsymbol{d}}_{\Lambda,R}'\dot{\boldsymbol{C}}_R\dot{\boldsymbol{d}}_{\beta,R}.\end{aligned}$$

It follows that

$$\frac{1}{T}\dot{\boldsymbol{d}}_{\beta,R}'\dot{\boldsymbol{A}}_R\dot{\boldsymbol{d}}_{\beta,R} - \frac{2}{T}\dot{\boldsymbol{d}}_{\Lambda,R}'\dot{\boldsymbol{C}}_R\dot{\boldsymbol{d}}_{\beta,R} + O_P\left(p^{1/2}\delta_{p,NT}^{-1}\right) \le 0.$$

This, in junction with the fact that

$$\left|\dot{\boldsymbol{d}}_{\Lambda,R}'\dot{\boldsymbol{C}}_R\dot{\boldsymbol{d}}_{\beta,R}\right| \le \left[\dot{\boldsymbol{d}}_{\Lambda,R}'\dot{\boldsymbol{d}}_{\Lambda,R}\right]^{1/2}\left[\dot{\boldsymbol{d}}_{\beta,R}'\dot{\boldsymbol{C}}_R'\dot{\boldsymbol{C}}_R\dot{\boldsymbol{d}}_{\beta,R}\right]^{1/2} \le \left\|\dot{\boldsymbol{d}}_{\Lambda,R}\right\|\left\|\dot{\boldsymbol{d}}_{\beta,R}\right\|\left[\mu_{\max}^{1/2}(\dot{\boldsymbol{C}}_R'\dot{\boldsymbol{C}}_R)\right],$$

implies that $\frac{1}{T}\dot{\boldsymbol{d}}_{\beta,R}'\dot{\boldsymbol{A}}_R\dot{\boldsymbol{d}}_{\beta,R} - \frac{2}{T^{1/2}}\left\|\dot{\boldsymbol{d}}_{\Lambda,R}\right\|\left\|\dot{\boldsymbol{d}}_{\beta,R}\right\|\mu_{\max}^{1/2}(\dot{\boldsymbol{C}}_R'\dot{\boldsymbol{C}}_R/T) + O_P(p^{1/2}\delta_{p,NT}^{-1}) \le 0$. Using a decomposition similar to (B.8) in Appendix B, we can readily show that $\mu_{\max}(\dot{\boldsymbol{C}}_R'\dot{\boldsymbol{C}}_R/T) = o_P(1)$. By Assumption 1(ii), $\mu_{\min}(\dot{\boldsymbol{A}}_R) > c_x$ w.p.a.1. and $||\dot{\boldsymbol{d}}_{\Lambda,R}|| = O_P(1)$. It follows that

$$\frac{1}{T}\|\dot{\boldsymbol{d}}_{\beta,R}\|^2 = \frac{1}{T}\sum_{t=1}^{T}\|\dot{\beta}_{t,R}-\beta_t^0\|^2 = o_P(1).$$

Note that $V(R,\dot{\boldsymbol{\beta}}_R) = \min_{\boldsymbol{\beta},\mathbf{\Lambda}_R}\hat{Q}_{NT}(\boldsymbol{\beta},\mathbf{\Lambda}_R)$ subject to $\mathbf{\Lambda}_R'\mathbf{\Lambda}_R/N = \boldsymbol{I}_R$. Let $s_r(\boldsymbol{\beta}) = \mu_r[\sum_{t=1}^{T}(Y_t - X_t\beta_t)(Y_t - X_t\beta_t)'/T]$. For any $R < R_0$, we make the following decomposition:

$$V(R,\boldsymbol{\beta}) = \frac{1}{N}\sum_{r=R_0+1}^{N}s_r(\boldsymbol{\beta}) + \frac{1}{N}\sum_{r=R+1}^{R_0}s_r(\boldsymbol{\beta}) \equiv S_1(\boldsymbol{\beta}) + S_{2R}(\boldsymbol{\beta}).$$

Noting that $S_1(\dot{\boldsymbol{\beta}}_R) \ge S_1(\dot{\boldsymbol{\beta}}_{R_0}) = V(R_0,\dot{\boldsymbol{\beta}}_{R_0})$, we have

$$V(R,\dot{\boldsymbol{\beta}}_R) - V(R_0,\dot{\boldsymbol{\beta}}_{R_0}) = \left[S_1(\dot{\boldsymbol{\beta}}_R) - S_1(\dot{\boldsymbol{\beta}}_{R_0})\right] + S_{2R}(\dot{\boldsymbol{\beta}}_R) \ge S_{2R}(\dot{\boldsymbol{\beta}}_R).$$

Let $s_r^0 = \mu_r\left(\frac{1}{T}\sum_{t=1}^T\left[\mathbf{\Lambda}^0 f_t^0 f_t^{0\prime}\mathbf{\Lambda}^{0\prime} + \varepsilon_t\varepsilon_t' + X_t(\beta_t^0 - \dot{\beta}_{t,R})(\beta_t^0 - \dot{\beta}_{t,R})'X_t'\right]\right)$. Notice that

$$
\begin{aligned}
& \frac{1}{N}\left|s_r(\dot{\boldsymbol{\beta}}_R) - s_r^0\right| \\
\leq\ & \frac{1}{NT}\Big\|\sum_{t=1}^T\Big\{(\mathbf{\Lambda}^0 f_t^0\varepsilon_t' + \varepsilon_t f_t^{0\prime}\mathbf{\Lambda}^{0\prime}) + [\mathbf{\Lambda}^0 f_t^0(\beta_t^0 - \dot{\beta}_{t,R})'X_t' + X_t(\beta_t^0 - \dot{\beta}_{t,R})f_t^{0\prime}\mathbf{\Lambda}^{0\prime}] \\
& + [\varepsilon_t(\beta_t^0 - \dot{\beta}_{t,R})'X_t' + X_t(\beta_t^0 - \dot{\beta}_{t,R})\varepsilon_t']\Big\}\Big\|_{\mathrm{sp}} \\
\leq\ & \frac{2}{NT}\left\|\sum_{t=1}^T\mathbf{\Lambda}^0 f_t^0\varepsilon_t'\right\|_{\mathrm{sp}} + \frac{2}{NT}\left\|\sum_{t=1}^T\mathbf{\Lambda}^0 f_t^0(\beta_t^0 - \dot{\beta}_{t,R})'X_t'\right\|_{\mathrm{sp}} + \frac{2}{NT}\left\|\sum_{t=1}^T\varepsilon_t(\beta_t^0 - \dot{\beta}_{t,R})'X_t'\right\|_{\mathrm{sp}}.
\end{aligned}
$$

Under Assumptions 1-2 and using the fact that $\frac{1}{T}\|\dot{\boldsymbol{d}}_{\beta,R}\|^2 = o_P(1)$, we can readily show that the second and third terms in the last expression are $o_P(1)$. The first term is $O_P((NT)^{-1/2})$ by Assumption 1(iii). It follows that

$$
\begin{aligned}
S_{2R}(\dot{\boldsymbol{\beta}}_R) &\geq \frac{1}{N}\sum_{r=R+1}^{R_0} s_r^0 + o_P(1) \\
&\geq \frac{1}{NT}\sum_{r=R+1}^{R_0}\mu_r\left(\mathbf{\Lambda}^0\mathbf{F}^{0\prime}\mathbf{F}^0\mathbf{\Lambda}^{0\prime}\right) + o_P(1) \\
&\geq (R_0 - R)\,\mu_{\min}(\mathbf{F}^{0\prime}\mathbf{F}^0/T)\mu_{\min}(\mathbf{\Lambda}^{0\prime}\mathbf{\Lambda}^0/N) + o_P(1) \\
&= (R_0 - R)\,\mu_{\min}(\mathbf{\Sigma}_F)\mu_{\min}(\mathbf{\Sigma}_\Lambda) + o_P(1),
\end{aligned}
$$

where the second inequality follows from Weyl's inequality. In sum, we have

$$
\operatorname{plim}\inf_{(N,T)\to\infty} V(R,\dot{\boldsymbol{\beta}}_R) - V(R_0,\dot{\boldsymbol{\beta}}_{R_0}) \geq c_R, \quad c_R = (R_0 - R)\,\mu_{\min}(\mathbf{\Sigma}_F)\mu_{\min}(\mathbf{\Sigma}_\Lambda)/2,
$$

completing the proof of Lemma B.4(i).

(ii) Recall that $V(R,\dot{\boldsymbol{\beta}}_R) = \min_{\boldsymbol{\beta},\mathbf{\Lambda}_R}\hat{Q}_{NT}(\boldsymbol{\beta},\mathbf{\Lambda}_R)$ subject to $\mathbf{\Lambda}_R'\mathbf{\Lambda}_R/N = \boldsymbol{I}_R$. Noting that $V(R,\dot{\boldsymbol{\beta}}_R) = \hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_R,\dot{\mathbf{\Lambda}}_R)$, by the triangle inequality, we have

$$
\begin{aligned}
& \left|V(R,\dot{\boldsymbol{\beta}}_R) - V(R_0,\dot{\boldsymbol{\beta}}_{R_0})\right| \\
\leq\ & \left|\hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_R,\dot{\mathbf{\Lambda}} - R) - \hat{Q}_{NT}(\boldsymbol{\beta}^0,\mathbf{\Lambda}^0)\right| + \left|\hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_{R_0},\dot{\mathbf{\Lambda}}_{R_0}) - \hat{Q}_{NT}(\boldsymbol{\beta}^0,\mathbf{\Lambda}^0)\right| \\
\leq\ & 2\max_{R_0\leq R\leq R_{\max}}\left|\hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_R,\dot{\mathbf{\Lambda}}_R) - \hat{Q}_{NT}(\boldsymbol{\beta}^0,\mathbf{\Lambda}^0)\right|.
\end{aligned}
$$

It suffices to show that $\hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_R,\dot{\mathbf{\Lambda}}_R) - \hat{Q}_{NT}(\boldsymbol{\beta}_0,\mathbf{\Lambda}_0) = O_P\left(\delta_{p,NT}^{-2}\right)$ for each $R \in [R_0, R_{\max}]$. Let $\dot{\boldsymbol{H}}_R^+$ denote the Moore-Penrose generalized inverse of $\dot{\boldsymbol{H}}_R$ such that $\dot{\boldsymbol{H}}_R\dot{\boldsymbol{H}}_R^+ = \boldsymbol{I}_{R_0}$; see, for example, the proof of Lemma C.6(iv). Noting that $Y_t - X_t\beta_t^0 = \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t$ and $\boldsymbol{M}_{\mathbf{\Lambda}^0}\mathbf{\Lambda}^0 = \mathbf{0}$, we may show that

$$\hat{Q}_{NT}(\boldsymbol{\beta}^0, \mathbf{\Lambda}^0) = \frac{1}{NT}\sum_{t=1}^{T}(Y_t - X_t\beta_t^0)'\boldsymbol{M}_{\mathbf{\Lambda}^0}(Y_t - X_t\beta_t^0) = \frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'\boldsymbol{M}_{\mathbf{\Lambda}^0}\varepsilon_t.$$

Let $\breve{\varepsilon}_t = \varepsilon_t - (\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^+ f_t^0$. Noting that

$$\begin{aligned} Y_t - X_t\dot{\beta}_{t,R} &= (X_t\beta_t^0 + \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t) - X_t\dot{\beta}_{t,R} \\ &= X_t(\beta_t^0 - \dot{\beta}_{t,R}) + \breve{\mathbf{\Lambda}}_R\dot{\boldsymbol{H}}_R^+ f_t^0 + \varepsilon_t + (\mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R - \breve{\mathbf{\Lambda}}_R)\dot{\boldsymbol{H}}_R^+ f_t^0 \\ &= X_t(\beta_t^0 - \dot{\beta}_{t,R}) + \breve{\mathbf{\Lambda}}_R\dot{\boldsymbol{H}}_R^+ f_t^0 + \breve{\varepsilon}_t \end{aligned}$$

and $\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\breve{\mathbf{\Lambda}}_R = \boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\left(\dot{\mathbf{\Lambda}}_R\dot{\boldsymbol{V}}_{NT,R}\right) = \mathbf{0}$, we have

$$\begin{aligned} \hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_R, \dot{\mathbf{\Lambda}}_R) &= \frac{1}{NT}\sum_{t=1}^{T}(Y_t - X_t\dot{\beta}_{t,R})'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}(Y_t - X_t\dot{\beta}_{t,R}) \\ &= \frac{1}{NT}\sum_{t=1}^{T}\left[X_t(\beta_t^0 - \dot{\beta}_{t,R}) + \breve{\varepsilon}_t\right]'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\left[X_t(\beta_t^0 - \dot{\beta}_{t,R}) + \breve{\varepsilon}_t\right] \\ &= \frac{1}{NT}\sum_{t=1}^{T}\breve{\varepsilon}_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\breve{\varepsilon}_t + \frac{1}{NT}\sum_{t=1}^{T}(\dot{\beta}_{t,R} - \beta_t^0)'X_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_t(\dot{\beta}_{t,R} - \beta_t^0) \\ &\quad - \frac{2}{NT}\sum_{t=1}^{T}\breve{\varepsilon}_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}X_t(\dot{\beta}_{t,R} - \beta_t^0) \\ &\equiv I_1 + I_2 - 2I_3. \end{aligned}$$

We next prove Lemma B.4(ii) by only showing that $I_1 - \hat{Q}_{NT}(\boldsymbol{\beta}^0, \mathbf{\Lambda}^0) = O_P(\delta_{p,NT}^{-2})$, $I_2 = O_P(\delta_{p,NT}^{-2})$, and $I_3 = O_P(\delta_{p,NT}^{-2})$.

First, using $\breve{\varepsilon}_t = \varepsilon_t - (\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^+ f_t^0$, we make the following decomposition:

$$\begin{aligned} I_1 &= \frac{1}{NT}\sum_{t=1}^{T}[\varepsilon_t - (\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^+ f_t^0]'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}[\varepsilon_t - (\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^+ f_t^0] \\ &= \frac{1}{NT}\sum_{r=1}^{T}\varepsilon_t'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\varepsilon_t - \frac{2}{NT}\sum_{r=1}^{T}f_t^{0\prime}\dot{\boldsymbol{H}}_R^{+\prime}(\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}\varepsilon_t \\ &\quad + \frac{1}{NT}\sum_{r=1}^{T}f_t^{0\prime}\dot{\boldsymbol{H}}_R^{+\prime}(\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)'\boldsymbol{M}_{\dot{\mathbf{\Lambda}}_R}(\breve{\mathbf{\Lambda}}_R - \mathbf{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^+ f_t^0 \\ &\equiv I_{1,1} - 2I_{1,2} + I_{1,3}. \end{aligned}$$

Using the arguments as in the proof of Lemmas C.1(iii)(iv), we can show that

$$I_{1,1} - \hat{Q}_{NT}(\boldsymbol{\beta}^0, \mathbf{\Lambda}^0) = \frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'(\boldsymbol{P}_{\mathbf{\Lambda}^0} - \boldsymbol{P}_{\dot{\mathbf{\Lambda}}^R})\varepsilon_t = O_P\left(\delta_{NT}^{-2}\right) = O_P\left(\delta_{p,NT}^{-2}\right).$$

For $I_{1,2}$, we have

$$\begin{aligned} I_{1,2} &= \frac{1}{NT}\sum_{t=1}^{T} f_t^{0\prime}\dot{\boldsymbol{H}}_R^{+\prime}\left(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R\right)'\varepsilon_t-\frac{1}{NT}\sum_{t=1}^{T} f_t^{0\prime}\dot{\boldsymbol{H}}_R^{+\prime}(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}_R}\varepsilon_t \\ &\equiv I_{1,2a}-I_{1,2b}. \end{aligned}$$

Using the decomposition in (C.110) and Lemma C.6(i), we can readily show that $I_{1,12a} = O_P\left(\delta_{p,NT}^{-2}\right)$. By the Cauchy-Schwarz inequality, the fact that $\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}_R}$ is a projection matrix, and Lemma C.1(iii),

$$\begin{aligned} |I_{1,2b}| &\leq \left[\frac{1}{NT}\sum_{t=1}^{T}\left\|(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^{+}f_t^0\right\|^2\right]^{1/2}\left[\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}_R}\varepsilon_t\right]^{1/2} \\ &= O_P\left(\delta_{p,NT}^{-1}\right)\cdot O_P\left(\delta_{NT}^{-1}\right)=O_P\left(\delta_{p,NT}^{-2}\right), \end{aligned}$$

where the following result which can be proved by Lemma C.6 has also been used:

$$\begin{aligned} \frac{1}{NT}\sum_{t=1}^{T}\left\|(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^{+}f_t^0\right\|^2 &\leq \frac{1}{N}\left\|\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R\right\|^2\left\|\dot{\boldsymbol{H}}_R^{+}\right\|^2\frac{1}{T}\sum_{t=1}^{T}\left\|f_t^0\right\|^2 \\ &= O_P\left(\delta_{p,NT}^{-2}\right). \end{aligned} \tag{C.112}$$

Thus we have $I_{1,2} = O_P\left(\delta_{p,NT}^{-2}\right)$. Similarly, using the fact that $\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}$ is a projection matrix and by (C.112), $I_{1,3} \leq \frac{1}{NT}\sum_{r=1}^{T}\left\|(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^{+}f_t^0\right\|^2 = O_P\left(\delta_{p,NT}^{-2}\right)$. As a consequence, we may complete the proof of $I_1-Q_{NT}(\boldsymbol{\beta}^0,\boldsymbol{\Lambda}^0)=O_P(\delta_{p,NT}^{-2})$ for each $R\in[R_0,R_{\max}]$.

Next, by Assumption 1(ii) and the fact that $\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}$ is a projection matrix and that $\dot{\eta}_R = \frac{1}{T}\sum_{t=1}^{T}||\dot{\beta}_{t,R}-\beta_t^0||^2 = O_P\left(\delta_{p,NT}^{-2}\right)$, we have

$$I_2 \leq \frac{1}{NT}\sum_{t=1}^{T}\left\|(\dot{\beta}_{t,R}-\beta_t^0)'X_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}X_t(\dot{\beta}_{t,R}-\beta_t^0)\right\| \leq \max_{1\leq t\leq T}\mu_{\max}\left(X_t'X_t/N\right)\dot{\eta}_R = O_P\left(\delta_{p,NT}^{-2}\right).$$

To study I_3, we apply $\breve{\varepsilon}_t=\varepsilon_t-(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^{+}f_t^0$ and $\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}=\boldsymbol{I}_N-\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}_R}$ and make the following decomposition:

$$\begin{aligned} I_3 &= \frac{1}{NT}\sum_{t=1}^{T}\breve{\varepsilon}_t'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}X_t(\dot{\beta}_{t,R}-\beta_t^0) \\ &= \frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'X_t(\dot{\beta}_{t,R}-\beta_t^0)-\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}_R}X_t(\dot{\beta}_{t,R}-\beta_t^0) \\ &\quad -\frac{1}{NT}\sum_{t=1}^{T}f_t^{0\prime}\dot{\boldsymbol{H}}_R^{+\prime}(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}X_t(\dot{\beta}_{t,R}-\beta_t^0) \\ &\equiv I_{3,1}-I_{3,2}-I_{3,3}. \end{aligned}$$

By the Cauchy-Schwarz inequality, Assumptions 1(ii)-(iii), the fact that

$$\dot{\eta}_R = \frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\beta}_{t,R}-\beta_t^0\right\|^2 = O_P\left(\delta_{p,NT}^{-2}\right), \quad \frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}_R}\varepsilon_t = O_P\left(\delta_{NT}^{-2}\right), \quad \mu_{\max}(\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}) = 1,$$

and Lemma C.6(i), we have

$$\begin{aligned}
|I_{3,1}| &\leq \left[\frac{1}{N^2T}\sum_{t=1}^{T}\varepsilon_t'X_tX_t'\varepsilon_t\right]^{1/2}\dot{\eta}_R^{1/2} = O_P(p^{1/2}N^{-1/2})O_P(\delta_{p,NT}^{-1}) = O_P\left(\delta_{p,NT}^{-2}\right), \\
|I_{3,2}| &\leq \left[\frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'\boldsymbol{P}_{\dot{\boldsymbol{\Lambda}}_R}\varepsilon_t\right]^{1/2}\left[\frac{1}{NT}\sum_{t=1}^{T}(\dot{\beta}_{t,R}-\beta_t^0)'X_t'X_t(\dot{\beta}_{t,R}-\beta_t^0)\right]^{1/2} \\
&\leq O_P\left(\delta_{NT}^{-1}\right)\mu_{\max}\left(X_t'X_t/N\right)^{1/2}\dot{\eta}_R^{1/2} = O_P\left(\delta_{p,NT}^{-2}\right),
\end{aligned}$$

and

$$\begin{aligned}
|I_{3,3}| &\leq \left[\frac{1}{NT}\sum_{t=1}^{T}f_t^{0\prime}\dot{\boldsymbol{H}}_R^{+\prime}(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)'\boldsymbol{M}_{\dot{\boldsymbol{\Lambda}}_R}(\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R)\dot{\boldsymbol{H}}_R^{+}f_t^0\right]^{1/2} \\
&\quad\times\left[\frac{1}{NT}\sum_{t=1}^{T}(\dot{\beta}_{t,R}-\beta_t^0)'X_t'X_t(\dot{\beta}_{t,R}-\beta_t^0)\right]^{1/2} \\
&\leq \frac{1}{N^{1/2}}\left\|\breve{\boldsymbol{\Lambda}}_R-\boldsymbol{\Lambda}^0\dot{\boldsymbol{H}}_R\right\|\left\|\dot{\boldsymbol{H}}_R^{+}\right\|\left[\frac{1}{T}\sum_{t=1}^{T}\left\|f_t^0\right\|^2\right]^{1/2}\mu_{\max}^{1/2}\left(X_t'X_t/N\right)\dot{\eta}_R^{1/2} \\
&= O_P(\delta_{p,NT}^{-1})O_P(1)\,O_P(\delta_{p,NT}^{-1}) = O_P\left(\delta_{p,NT}^{-2}\right).
\end{aligned}$$

Hence $I_3 = O_P\left(\delta_{p,NT}^{-2}\right)$. In sum, we have shown that $\hat{Q}_{NT}(\dot{\boldsymbol{\beta}}_R,\dot{\boldsymbol{\Lambda}}_R)-\hat{Q}_{NT}\left(\boldsymbol{\beta}_0,\boldsymbol{\Lambda}_0\right) = O_P\left(\delta_{p,NT}^{-2}\right)$ for each $R\in[R_0,R_{\max}]$, completing the proof of Lemma B.4(ii). ■

Proof of Lemma B.5. Let

$$D_{NT}\left(\boldsymbol{\alpha}_m,\boldsymbol{\Lambda};\mathcal{T}_m\right) = \frac{1}{NT}\sum_{j=1}^{m+1}\sum_{t=T_{j-1}}^{T_j-1}\left[(Y_t-X_t\alpha_j)'\boldsymbol{M}_{\boldsymbol{\Lambda}}\left(Y_t-X_t\alpha_j\right)-\varepsilon_t'\varepsilon_t\right]$$

and $\bar{\sigma}_{NT}^2 = \frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'\varepsilon_t$. Note that

$$\left(\tilde{\boldsymbol{\alpha}}_m(\mathcal{T}_m),\tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m)\right) = \arg\min_{(\boldsymbol{\alpha}_m,\boldsymbol{\Lambda})}D_{NT}\left(\boldsymbol{\alpha}_m,\boldsymbol{\Lambda};\mathcal{T}_m\right),$$

and

$$\tilde{\sigma}^2(\mathcal{T}_m)-\tilde{\sigma}^2(\mathcal{T}_{m^0}^0) = \left[\tilde{\sigma}^2(\mathcal{T}_m)-\bar{\sigma}_{NT}^2\right]-\left[\tilde{\sigma}^2(\mathcal{T}_{m^0}^0)-\bar{\sigma}_{NT}^2\right]$$

with $\tilde{\sigma}^2(\mathcal{T}_m) - \bar{\sigma}^2_{NT} = D_{NT}(\tilde{\boldsymbol{\alpha}}_m(\mathcal{T}_m), \tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m); \mathcal{T}_m)$. We prove the lemma by showing that (i)

$$\frac{m^0}{T\Delta^2_{NT}} \left[\tilde{\sigma}^2(\mathcal{T}^0_{m^0}) - \bar{\sigma}^2_{NT}\right] = o_P(1); \tag{C.113}$$

and (ii)

$$\frac{m^0}{T\Delta^2_{NT}} (\tilde{\sigma}^2(\mathcal{T}_m) - \bar{\sigma}^2_{NT}) \geq c + o_P(1) \ \text{ w.p.a.1 } \text{ for } \text{ some } \ c > 0. \tag{C.114}$$

We first show (C.113) in (i). We make the following decomposition:

$$\begin{aligned}
\tilde{\sigma}^2_{\mathcal{T}^0_{m^0}} &= \frac{1}{NT} \sum_{j=1}^{m^0+1} \sum_{t=T^0_{j-1}}^{T^0_j - 1} [Y_t - X_t \tilde{\alpha}_j]' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} [Y_t - X_t \tilde{\alpha}_j] \\
&= \frac{1}{NT} \sum_{j=1}^{m^0+1} \sum_{t=T^0_{j-1}}^{T^0_j - 1} \left[X_t(\alpha^0_j - \tilde{\alpha}_j) + \boldsymbol{\Lambda}^0 f^0_t + \varepsilon_t\right]' \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \left[X_t(\alpha^0_j - \tilde{\alpha}_j) + \boldsymbol{\Lambda}^0 f^0_t + \varepsilon_t\right] \\
&= \frac{1}{NT} \sum_{j=1}^{m^0+1} \sum_{t=T^0_{j-1}}^{T^0_j - 1} \left[\varepsilon'_t \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \varepsilon_t + f^{0\prime}_t \boldsymbol{\Lambda}^{0\prime} \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \boldsymbol{\Lambda}^0 f^0_t + (\alpha^0_j - \tilde{\alpha}_j)' X'_t \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} X_t (\alpha^0_j - \tilde{\alpha}_j)\right. \\
&\quad \left. + 2\varepsilon'_t \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} X_t(\alpha^0_j - \tilde{\alpha}_j) + 2\varepsilon'_t \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \boldsymbol{\Lambda}^0 f^0_t + 2 f^{0\prime}_t \boldsymbol{\Lambda}^{0\prime} \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} X_t(\alpha^0_j - \tilde{\alpha}_j)\right] \\
&\equiv d_{1NT} + d_{2NT} + d_{3NT} + 2d_{4NT} + 2d_{5NT} + 2d_{6NT},
\end{aligned}$$

where we suppress the dependence of $\tilde{\alpha}_j = \tilde{\alpha}_j(\mathcal{T}^0_{m^0})$ and $\tilde{\boldsymbol{\Lambda}} = \tilde{\boldsymbol{\Lambda}}(\mathcal{T}^0_{m^0})$ on $\mathcal{T}^0_{m^0}$ for notational simplicity. By Lemma C.1(iii),

$$d_{1NT} = \frac{1}{NT} \sum_{t=1}^{T} \varepsilon'_t \boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} \varepsilon_t = \frac{1}{NT} \sum_{t=1}^{T} \varepsilon'_t \varepsilon_t + O_P\left(\delta^{-2}_{NT}\right) = \bar{\sigma}^2_{NT} + O_P\left(\delta^{-2}_{NT}\right).$$

Using the preliminary results in Lemmas C.4 and C.5(i) and Theorem 3.4, we may show that $d_{lNT} = O_P(\delta^{-2}_{p,NT})$ for $l = 3, 4, 6$. Using $\boldsymbol{M}_{\boldsymbol{\Lambda}^0} \boldsymbol{\Lambda}^0 = 0$ and (C.79), and decomposing $\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} - \boldsymbol{M}_{\boldsymbol{\Lambda}^0} = -(\boldsymbol{P}_{\tilde{\boldsymbol{\Lambda}}} - \boldsymbol{P}_{\boldsymbol{\Lambda}^0})$ as in (C.81), we can readily show that

$$\begin{aligned}
d_{2NT} &= \frac{1}{NT} \sum_{t=1}^{T} f^{0\prime}_t \boldsymbol{\Lambda}^{0\prime} \left(\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} - \boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \boldsymbol{\Lambda}^0 f^0_t = O_P\left(\delta^{-2}_{p,NT}\right), \text{ and} \\
d_{5NT} &= \frac{1}{NT} \sum_{t=1}^{T} \varepsilon'_t \left(\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}} - \boldsymbol{M}_{\boldsymbol{\Lambda}^0}\right) \boldsymbol{\Lambda}^0 f^0_t = O_P\left(\delta^{-2}_{p,NT}\right).
\end{aligned}$$

It follows that

$$\tilde{\sigma}^2(\mathcal{T}^0_{m^0}) - \bar{\sigma}^2_{NT} = O_P\left(\delta^{-2}_{p,NT}\right), \tag{C.115}$$

which, together with Assumption 2(ii), leads to (C.113).

We now show (C.114) in (ii). We consider three cases: (a) $m^0 = 1$, (b) $m^0 = 2$, and (c) $3 < m^0 \leq m_{\max}$. For case (a) of $m^0 = 1$, if $n < m^0$, we have $m = 0$ and $\mathcal{T}_m = \mathcal{T}_0 = \varnothing$. The true model contains one structural break:

$$Y_t = \begin{cases} X_t \alpha_1^0 + \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t & \text{if } 1 \leq t \leq T_1^0 - 1, \\ X_t \alpha_2^0 + \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t & \text{if } T_1^0 \leq t \leq T; \end{cases}$$

while the working model that ignores the structural break in the regression coefficient is

$$Y_t = X_t \alpha + \mathbf{\Lambda}^0 f_t^0 + e_t, \ 1 \leq t \leq T,$$

where e_t is the error term. Note that $\tilde{\sigma}^2(\mathcal{T}_0) = \frac{1}{NT}\sum_{t=1}^T (Y_t - X_t\tilde{\alpha})' \boldsymbol{M}_{\tilde{\mathbf{\Lambda}}} (Y_t - X_t\tilde{\alpha})$, where

$$(\tilde{\alpha}, \tilde{\mathbf{\Lambda}}) = \arg\min_{\alpha, \mathbf{\Lambda}} \frac{1}{NT} \sum_{t=1}^T (Y_t - X_t\alpha)' \boldsymbol{M}_{\mathbf{\Lambda}} (Y_t - X_t\alpha)$$

subject to $\mathbf{\Lambda}'\mathbf{\Lambda}/N = \boldsymbol{I}_{R_0}$, and we suppress the dependence of $\tilde{\alpha}$ and $\tilde{\mathbf{\Lambda}}$ on $\mathcal{T}_0$. Using $Y_t - X_t\alpha = X_t(\beta_t^0 - \alpha) + \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t$ and Lemmas C.1(i)(ii), we can readily show that

$$\begin{aligned}
& \frac{1}{NT} \sum_{t=1}^T (Y_t - X_t\alpha)' \boldsymbol{M}_{\mathbf{\Lambda}} (Y_t - X_t\alpha) \\
= \ & \frac{1}{NT} \sum_{t=1}^T \left[X_t(\beta_t^0 - \alpha) + \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t\right]' \boldsymbol{M}_{\mathbf{\Lambda}} \left[X_t(\beta_t^0 - \alpha) + \mathbf{\Lambda}^0 f_t^0 + \varepsilon_t\right] \\
= \ & \frac{1}{NT} \sum_{t=1}^T \left[X_t(\beta_t^0 - \alpha) + \mathbf{\Lambda}^0 f_t^0\right]' \boldsymbol{M}_{\mathbf{\Lambda}} \left[X_t(\beta_t^0 - \alpha) + \mathbf{\Lambda}^0 f_t^0\right] + \frac{1}{NT}\sum_{t=1}^T \varepsilon_t'\varepsilon_t + O_P(p^{1/2}\delta_{p,NT}^{-1})
\end{aligned}$$

uniformly in α and $\mathbf{\Lambda}$ such that $\mathbf{\Lambda}'\mathbf{\Lambda}/N = \boldsymbol{I}_{R_0}$ and $\|\alpha\| \leq Cp^{1/2}$. It follows that

$$\begin{aligned}
\tilde{\sigma}^2(\mathcal{T}_0) &= \frac{1}{NT}\sum_{t=1}^T \tilde{Y}_t' \boldsymbol{M}_{\tilde{\mathbf{\Lambda}}} \tilde{Y}_t + \bar{\sigma}_{NT}^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}) \\
&\geq \min_{\mathbf{\Lambda}:\ \mathbf{\Lambda}'\mathbf{\Lambda}/N = \boldsymbol{I}_{R_0}} \frac{1}{NT}\sum_{t=1}^T \tilde{Y}_t' \boldsymbol{M}_{\mathbf{\Lambda}} \tilde{Y}_t + \bar{\sigma}_{NT}^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}) \\
&= \frac{1}{NT}\sum_{r=R_0+1}^N \mu_r \left[\sum_{t=1}^T \tilde{Y}_t\tilde{Y}_t'\right] + \bar{\sigma}_{NT}^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}) \\
&\geq \frac{1}{NT}\sum_{r=R_0+1}^N \mu_r \left[\sum_{t=1}^T X_t(\beta_t^0 - \tilde{\alpha})(\beta_t^0 - \tilde{\alpha})' X_t'\right] + \bar{\sigma}_{NT}^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}) \\
&= \frac{1}{NT}\min_{\mathbf{\Lambda}:\ \mathbf{\Lambda}'\mathbf{\Lambda}/N = \boldsymbol{I}_{R_0}} \left[\sum_{t=1}^T (\beta_t^0 - \tilde{\alpha})' X_t' \boldsymbol{M}_{\mathbf{\Lambda}} X_t (\beta_t^0 - \tilde{\alpha})\right] + \bar{\sigma}_{NT}^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}) \\
&\geq c_x \cdot \frac{1}{T}\sum_{t=1}^T \left\|\beta_t^0 - \tilde{\alpha}\right\|^2 + \bar{\sigma}_{NT}^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}),
\end{aligned}$$

where $\tilde{Y}_t = X_t(\beta_t^0 - \tilde{\alpha}) + \mathbf{\Lambda}^0 f_t^0$, the second and third inequalities follow from Weyl's inequality and Assumption 1(ii), respectively. Consequently, we have by Assumptions 5(i)-(ii)

$$\frac{m^0}{T\Delta_{NT}^2}\left[\tilde{\sigma}^2(\mathcal{T}_0) - \bar{\sigma}_{NT}^2\right] \geq c_x c_\beta + o_P(1),$$

where c_β is defined in Assumption 5(i). We have completed the proof of (C.114) for case (a).

In cases (b)-(c), it suffices to consider the case where $m = m^0 - 1$ (If $m < m^0 - 1$, one can always augment the set $\mathcal{T}_m$ by $m^0 - 1 - m$ true break points which are not inside $\mathcal{T}_m$ to make $D_{NT}(\tilde{\boldsymbol{\alpha}}_m(\mathcal{T}_m), \tilde{\mathbf{\Lambda}}(\mathcal{T}_m); \mathcal{T}_m)$ smaller). For the case (b) with $m = 1$, we consider three subcases: (b.1) $2 \leq T_1 \leq T_1^0$, (b.2) $T_1^0 < T_1 \leq T_2^0$, and (b.3) $T_2^0 < T_1 \leq T$. In the subcase (b.1), $[1, T_1 - 1]$ does not contain a break point while $[T_1, T]$ contains two true break points T_1^0 and T_2^0. Observe that

$$\begin{aligned}
D_{NT}(\tilde{\boldsymbol{\alpha}}_1(\mathcal{T}_1), \tilde{\mathbf{\Lambda}}(\mathcal{T}_1); \mathcal{T}_1) &= \frac{1}{NT}\sum_{t=1}^{T_1-1}\left\{[Y_t - X_t\tilde{\alpha}_1(\mathcal{T}_1)]'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_1)}[Y_t - X_t\tilde{\alpha}_1(\mathcal{T}_1)] - \varepsilon_t'\varepsilon_t\right\} \\
&\quad + \frac{1}{NT}\sum_{t=T_1}^{T}\left\{[Y_t - X_t\tilde{\alpha}_2(\mathcal{T}_1)]'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_1)}[Y_t - X_t\tilde{\alpha}_2(\mathcal{T}_1)] - \varepsilon_t'\varepsilon_t\right\} \\
&\equiv D_{NT,1} + D_{NT,2}.
\end{aligned}$$

Noting that the interval $[1, T_1 - 1]$ does not contain a break point, using the arguments as used in the study of case (a), we can readily show that

$$D_{NT,1} \geq \frac{c_x}{T}\sum_{t=1}^{T_1-1}\left\|\alpha_1^0 - \tilde{\alpha}_1(\mathcal{T}_1)\right\|^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}).$$

Similarly, we can show that

$$D_{NT,2} \geq \frac{c_x}{T}\sum_{t=T_1}^{T}\left\|\beta_t^0 - \tilde{\alpha}_2(\mathcal{T}_1)\right\|^2 + O_P(p^{1/2}\delta_{p,NT}^{-1}).$$

Then by Assumptions 5(i)(ii)

$$\begin{aligned}
&\frac{m^0}{T\Delta_{NT}^2}D_{NT}(\tilde{\boldsymbol{\alpha}}_1(\mathcal{T}_1), \tilde{\mathbf{\Lambda}}(\mathcal{T}_1); \mathcal{T}_1) \\
&\geq \frac{m^0}{T\Delta_{NT}^2}\left\{\frac{c_x}{T}\sum_{t=1}^{T_1-1}\left\|\alpha_1^0 - \tilde{\alpha}_1(\mathcal{T}_1)\right\|^2 + \frac{c_x}{T}\sum_{t=T_1}^{T}\left\|\beta_t^0 - \tilde{\alpha}_2(\mathcal{T}_1)\right\|^2 + O_P(p^{1/2}\delta_{p,NT}^{-1})\right\} \\
&\geq c_x \min_{\alpha_1,\alpha_2}\frac{m^0}{T\Delta_{NT}^2}\sum_{j=1}^{2}\sum_{t=T_{j-1}}^{T_j-1}\left\|\beta_t^0 - \alpha_j\right\|^2 + o_P(1) \\
&\geq c_x c_\beta + o_P(1).
\end{aligned}$$

In the subcase (b.2), both $[2,T_1-1]$ and $[T_1,T]$ contain a break. As in subcase (b.1), we can show that

$$\begin{aligned}
& \frac{m^0}{T\Delta_{NT}^2} D_{NT}(\tilde{\boldsymbol{\alpha}}_1(\mathcal{T}_1), \tilde{\boldsymbol{\Lambda}}(\mathcal{T}_1);\mathcal{T}_1) \\
\geq\ & \frac{m^0}{T\Delta_{NT}^2}\left\{\frac{c_x}{T}\sum_{t=1}^{T_1-1}\left\|\beta_t^0-\tilde{\alpha}_1(\mathcal{T}_1)\right\|^2+\frac{c_x}{T}\sum_{t=T_1}^{T}\left\|\beta_t^0-\tilde{\alpha}_2(\mathcal{T}_1)\right\|^2+O_P(pN^{-1/2}+p^{1/2}T^{-1/2})\right\} \\
\geq\ & c_x\min_{\alpha_1,\alpha_2}\frac{m^0}{T\Delta_{NT}^2}\sum_{j=1}^{2}\sum_{t=T_{j-1}}^{T_j-1}\left\|\beta_t^0-\alpha_j\right\|^2 \geq c_x c_\beta + o_P(1).
\end{aligned}$$

The proof for the subcase (b.3) is analogous to that for the subcase (b.1). Hence, the conclusion (C.114) follows in the subcase (b). Case (c) can be studied analogously. This completes the proof of the lemma. ■

Proof of Lemma B.6. For $\mathcal{T}_m \in \bar{\mathbb{T}}_m$ with $m^0 < m \leq m_{\max}$, we recall that

$$\begin{aligned}
\tilde{\sigma}^2(\mathcal{T}_m) &= Q_{NT}(\tilde{\boldsymbol{\alpha}}_m(\mathcal{T}_m), \tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m);\mathcal{T}_m) \\
&= \min_{\boldsymbol{\alpha}_m,\boldsymbol{\Lambda}}\frac{1}{NT}\sum_{j=1}^{m+1}\sum_{t=T_{j-1}}^{T_j-1}(Y_t-X_t\alpha_j)'\boldsymbol{M}_{\boldsymbol{\Lambda}}(Y_t-X_t\alpha_j) \\
&= \min_{\boldsymbol{\alpha}_m}\frac{1}{NT}\sum_{j=1}^{m+1}\sum_{t=T_{j-1}}^{T_j-1}(Y_t-X_t\alpha_j)'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m)}(Y_t-X_t\alpha_j),
\end{aligned}$$

and $\bar{\sigma}_{NT}^2 = \frac{1}{NT}\sum_{t=1}^{T}\varepsilon_t'\varepsilon_t$. In view of the fact that

$$\tilde{\sigma}^2(\mathcal{T}_{m^0}^0) \geq \tilde{\sigma}^2(\mathcal{T}_m) \ \text{ and } \ \tilde{\sigma}^2(\mathcal{T}_{m^0}^0) = \bar{\sigma}_{NT}^2 + O_P(\delta_{p,NT}^{-2})$$

by (C.115), we have

$$0 \leq \tilde{\sigma}^2(\mathcal{T}_{m^0}^0) - \tilde{\sigma}^2(\mathcal{T}_m) = \bar{\sigma}_{NT}^2 - \tilde{\sigma}^2(\mathcal{T}_m) + O_P(\delta_{p,NT}^{-2}) = \sum_{j=1}^{m+1} J_{NT,j} + O_P(\delta_{p,NT}^{-2}), \qquad \text{(C.116)}$$

where $J_{NT,j} \equiv -\inf_\alpha S_j(\alpha)$, $S_j(\alpha) = \frac{1}{NT}\sum_{t=T_{j-1}}^{T_j-1}\left[(Y_t-X_t\alpha)'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m)}(Y_t-X_t\alpha)-\varepsilon_t'\varepsilon_t\right]$ and $[T_{j-1},T_j-1]$ does not contain any break point for $j=1,...,m+1$. Let $\alpha_{j,m}^0 = \beta_{T_{j-1}}^0$ and $\tilde{\alpha}_{j,m} = \tilde{\alpha}_j(\mathcal{T}_m) = \arg\min_\alpha S_j(\alpha) = \left(\sum_{t=T_{j-1}}^{T_j-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m)}X_t\right)^{-1}\sum_{t=T_{j-1}}^{T_j-1}X_t'\boldsymbol{M}_{\tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m)}Y_t$ for $j=1,...,m+1$. As in the proofs of Lemma C.4(i) and Theorems 3.1 and 3.4, we can show that $\frac{1}{N}||\tilde{\boldsymbol{\Lambda}}(\mathcal{T}_m)-\boldsymbol{\Lambda}^0||^2 = O_P(\delta_{p,NT}^{-2})$ and $||\tilde{\alpha}_{j,m}-\alpha_{j,m}^0|| = O_P(\delta_{p,NT}^{-1})$. Then using $Y_t - X_t\tilde{\alpha}_{j,m} =$

$\varepsilon_t + \mathbf{\Lambda}^0 f_t^0 + X_t(\alpha_{j,m}^0 - \tilde{\alpha}_{j,m})$, we have

$$
\begin{aligned}
S_j\left(\tilde{\alpha}_{j,m}\right) &= \frac{1}{NT}\sum_{t=T_{j-1}}^{T_j-1}\left[\left(Y_t - X_t\tilde{\alpha}_{j,m}\right)'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)}\left(Y_t - X_t\tilde{\alpha}_{j,m}\right) - \varepsilon_t'\varepsilon_t\right] \\
&= \frac{1}{NT}\sum_{t=T_{j-1}}^{T_j-1}\left\{\left[\varepsilon_t + \mathbf{\Lambda}^0 f_t^0 + X_t(\alpha_{j,m}^0 - \tilde{\alpha}_{j,m})\right]'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)}\left[\varepsilon_t + \mathbf{\Lambda}^0 f_t^0 + X_t(\alpha_{j,m}^0 - \tilde{\alpha}_{j,m})\right] - \varepsilon_t'\varepsilon_t\right\} \\
&= \frac{-1}{NT}\sum_{t=T_{j-1}}^{T_j-1}\varepsilon_t'\boldsymbol{P}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)}\varepsilon_t + \frac{1}{NT}\sum_{t=T_{j-1}}^{T_j-1} f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)}\mathbf{\Lambda}^0 f_t^0 \\
&\quad + \frac{1}{NT}\sum_{t=T_{j-1}}^{T_j-1}\left(\alpha_{j,m}^0 - \tilde{\alpha}_{j,m}\right)' X_t'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)} X_t\left(\alpha_{j,m}^0 - \tilde{\alpha}_{j,m}\right) + \frac{2}{NT}\sum_{t=T_{j-1}}^{T_j-1}\varepsilon_t'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)}\mathbf{\Lambda}^0 f_t^0 \\
&\quad + \frac{2}{NT}\sum_{t=T_{j-1}}^{T_j-1}\varepsilon_t'\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)} X_t\left(\alpha_{j,m}^0 - \tilde{\alpha}_{j,m}\right) + \frac{2}{NT}\sum_{t=T_{j-1}}^{T_j-1} f_t^{0\prime}\mathbf{\Lambda}^{0\prime}\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)} X_t\left(\alpha_{j,m}^0 - \tilde{\alpha}_{j,m}\right) \\
&\equiv S_{j,1} + S_{j,2} + S_{j,3} + 2S_{j,4} + 2S_{j,5} + 2S_{j,6}.
\end{aligned}
$$

By Lemma C.1(iii),

$$
\sum_{j=1}^{m+1} S_{j,1} = \frac{-1}{NT}\sum_{t=1}^{T}\varepsilon_t'\boldsymbol{P}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)}\varepsilon_t = O_P\left(\delta_{NT}^{-2}\right).
$$

In addition, we can show that

$$
\begin{aligned}
\sum_{j=1}^{m+1} S_{j,2} &= \frac{1}{NT}\sum_{t=1}^{T} f_t^{0\prime}\mathbf{\Lambda}^{0\prime}(\boldsymbol{M}_{\tilde{\mathbf{\Lambda}}(\mathcal{T}_m)} - \boldsymbol{M}_{\mathbf{\Lambda}^0})\mathbf{\Lambda}^0 f_t^0 = O_P\left(\delta_{p,NT}^{-2}\right), \\
\sum_{j=1}^{m+1} S_{j,3} &\leq \frac{1}{T}\sum_{j=1}^{m+1}\left\|\alpha_{j,m}^0 - \tilde{\alpha}_{j,m}\right\|^2\sum_{t=T_{j-1}}^{T_j-1}\mu_{\max}\left(X_t'X_t/N\right) = O_P\left(\delta_{p,NT}^{-2}\right),
\end{aligned}
$$

and similarly $\sum_{j=1}^{m+1} S_{j,l} = O_P\left(\delta_{p,NT}^{-2}\right)$ for $l = 4, 5, 6$. Then by (C.116), $\tilde{\sigma}^2(\mathcal{T}_m) - \bar{\sigma}_{NT}^2 = O_P\left(\delta_{p,NT}^{-2}\right)$ for all $m \in \left\{m^0+1, ..., m_{\max}\right\}$ and $\mathcal{T}_m = \{T_1, ..., T_m\}$, which completes the proof of Lemma B.6. ■

References

Bai, J. (2003), "Inferential theory for factor models with large dimensions," *Econometrica,* 71, 135-171.

Bai, J. (2009), "Panel data models with interactive fixed effects," *Econometrica,* 77, 1229-1279.

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

Bernstein, D. S. (2005), Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory. Princeton University Press, Princeton.