IMPUTED FACTOR REGRESSION FOR HIGH-DIMENSIONAL BLOCK-WISE MISSING DATA

Yanqing Zhang, Niansheng Tang and Annie Qu

Yunnan University and University of Illinois at Urbana-Champaign

Supplementary Material

S1 Assumptions for the theoretical properties of the proposed method

Let $\|\mathbf{A}\| = \{\operatorname{tr}(\mathbf{A}^\top\mathbf{A})\}^{1/2}$ be the norm of matrix $\mathbf{A}$ and C denote a positive constant. The following assumptions are needed to establish the theoretical properties in Section 3.3.

(C1) Factors: $E(\|\mathbf{w}_t\|^4) \leqslant C < \infty$, $n_1^{-1}\sum_{t=1}^{n_1}\mathbf{W}_{(1)t}\mathbf{W}_{(1)t}^\top \xrightarrow{\text{p}} \boldsymbol{\Sigma}_{\text{w}}$ and $n^{-1}\sum_{t=1}^{n}\mathbf{w}_t\mathbf{w}_t^\top \xrightarrow{\text{p}} \boldsymbol{\Sigma}_{\text{w}}$ for some $r \times r$ positive definite matrix $\boldsymbol{\Sigma}_{\text{w}}$, where $\mathbf{W}_{(1)t}$ is an $r \times 1$ vector from the tth row of $\mathbf{W}_{(1)} \in \mathbb{R}^{n_1 \times r}$.

(C2) Factor loadings: $\|\boldsymbol{\lambda}_i\| \leqslant \bar{\lambda} < \infty$ and $\|\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}/q - \boldsymbol{\Sigma}_\Lambda\| \to 0$ for some positive finite value $\bar{\lambda}$ and some $r \times r$ positive definite matrix $\boldsymbol{\Sigma}_\Lambda$.

(C3) Weak dependence: (i) $E(e_{ti}) = 0$, $E(|e_{ti}|^8) \leqslant C$; (ii) $E(e_{ti}e_{lj}) = \tau_{ij}$

if $t = l$ and 0 otherwise, $q^{-1}\sum_{i=1}^{q}\tau_{ii} \leqslant C$, $q^{-1}\sum_{i=1}^{q}\sum_{j=1}^{q}|\tau_{ij}| \leqslant C$, $\sum_{j=1}^{q}|\tau_{ij}| \leqslant C$, $q^{-1}\sum_{i=1}^{q}\sum_{j=1}^{q}\tau_{ij}^2 \leqslant C$; (iii) $q^{-1}\sum_{i=1}^{q}\sum_{j=1}^{q}\{E(e_{ti}^2e_{tj}^2)-\tau_{ij}^2\} \leqslant C$; (iv) $E(|q^{-1/2}\sum_{i=1}^{q}\{e_{ti}e_{li} - E(e_{ti}e_{li})\}|^4) \leqslant C$ for every (t, l).

(C4) $E\{q^{-1}\sum_{i=1}^{q}\|n_1^{-1/2}\sum_{t=1}^{n_1}\mathbf{W}_{(1)t}\boldsymbol{e}_{(1)ti}\|^2\} \leqslant C$.

(C5) For any $t = 1, \ldots, n_1$, $E(\|(n_1q)^{-1/2}\sum_{i=1}^{q}\sum_{t=1}^{n_1}\mathbf{W}_{(1)t}\boldsymbol{\lambda}_i^\top\boldsymbol{e}_{(1)ti}\|^2) \leqslant C$ and $E(\|(n_1q)^{-1/2}\sum_{i=1}^{q}\sum_{l=1}^{n_1}\mathbf{W}_{(1)l}[\boldsymbol{e}_{(1)li}\boldsymbol{e}_{(1)ti} - E\{\boldsymbol{e}_{(1)li}\boldsymbol{e}_{(1)ti}\}]\|^2) \leqslant C$.

(C6) Predicting model: $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) < C < \infty$, $n^{-1}\sum_t^n \mathbf{w}_t\varepsilon_t \xrightarrow{\text{p}} 0$ and $\|\boldsymbol{\alpha}\| < \infty$.

Assumption (C1) is general for the factor model where components of factor variables are correlated. Assumption (C2) ensures that each factor has a nontrivial contribution to the variance of $\boldsymbol{z}_t$. Here we only consider non-random factor loadings for simplicity. Assumption (C3) is the weak correlation assumption. Given Assumption (C3)(i), the remaining assumptions in (C3) are satisfied if the e_{ti}'s are independent for all i. Assumptions (C3)(iii) and (iv) imply that the fourth and the eighth moments are bounded, respectively. Thus, the proposed method is applicable for sub-Gaussian cases. The assumption that $\sum_{j=1}^{q}|\tau_{ij}| \leqslant C$ for all i in Assumption (C3)(ii) implies that the eigenvalues of the covariance matrix of random error $\boldsymbol{e}_i$ are bounded, since the largest eigenvalue of the covariance matrix is bounded by

$\max_i \sum_{j=1}^{q} |\tau_{ij}|$. Assumption (C4) provides weak dependence between factors and random errors. When factors and errors are independent, which is a standard assumption for conventional factor models, Assumption (C4) is implied by Assumptions (C1) and (C3), although independence is not required for Assumption (C4) to hold. Assumption (C5) is not restrictive since the sums in Assumption (C5) involve zero mean random variables. Assumption (C6) is a standard set of conditions that implies consistency of the ordinary least square estimator in the predicting model.

S2 The accuracy of screening for high-dimensional block-wise missing data

Let $\mathcal{M}_o$ be the collection of indices for nonzero parameters in true sparse model $y = \sum_{\ell=1}^{p} X_\ell \beta_\ell + \varepsilon$, and p_o be the number of elements in $\mathcal{M}_o$. Define $\mathbf{\Sigma} = \mathrm{cov}(\boldsymbol{x})$ and $\widetilde{\mathbf{X}}_o^{(i)} = \mathbf{X}_o^{(i)} \mathbf{\Sigma}_{(i)}^{-1/2}$ for $i = 1, \ldots, K$, where $\mathbf{X}_o^{(i)}$ is an $n_{o(i)} \times s_i$ matrix of observed values from the ith data source and $\mathbf{\Sigma}_{(i)}$ is the corresponding submatrix of $\mathbf{\Sigma}$. Thus, the covariance matrix of the transformed matrix $\widetilde{\mathbf{X}}_o^{(i)}$ is $\mathbf{I}_{s_i}$. The following assumptions are needed for the accuracy of screening.

(A1) $\mathrm{Var}(y) = O(1)$, and for some $\kappa \geqslant 0$ and $c_1, c_2 > 0$,

$$\min_{\ell \in \mathcal{M}_o} |\beta_\ell| \geqslant \frac{c_1}{\min_{1 \leqslant i \leqslant K} n_{o(i)}^{\kappa}} \qquad \text{and} \qquad \min_{\ell \in \mathcal{M}_o} |\mathrm{cov}(\beta_\ell^{-1} y, X_\ell)| \geqslant c_2.$$

(A2) For $i = 1, \ldots, K$, $\widetilde{\mathbf{X}}_o^{(i)}$ has a continuous and spherically symmetric distribution. If there are some $c_3 > 1$ and $C_1 > 0$ such that the deviation inequality $\Pr\{\lambda_{\max}(s_i^{-1} \widetilde{\mathbf{X}}_o^{(i)} \widetilde{\mathbf{X}}_o^{(i)\top}) > c_3 \text{ or } \lambda_{\min}(s_i^{-1} \widetilde{\mathbf{X}}_o^{(i)} \widetilde{\mathbf{X}}_o^{(i)\top}) < 1/c_3\} \leqslant \exp(-C_1 n_{o(i)})$ holds, where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ are the largest and smallest eigenvalues of a matrix, respectively. Also, $\varepsilon \sim N(0, \sigma^2)$ for some $\sigma > 0$.

(A3) There are some $\tau \geqslant 0$ and $c_4 > 0$ such that $\lambda_{\max}(\mathbf{\Sigma}_{(i)}) \leqslant c_4 n_{o(i)}^{\tau}$ for $i = 1, \ldots, K$.

(A4) $p > n$ and $\log(p) = O(n^{\varrho})$ for some $\varrho \in (0, 1 - 2\kappa)$, where κ is given by condition (A1).

The following lemma gives the accuracy of screening for high-dimensional block-wise missing data.

Lemma 1. (accuracy of screening) *Under Conditions (A1)-(A4), if* $2\kappa + \tau < 1$ *then there is some* $\gamma \in (0, 1)$ *such that, when* $\gamma \to 0$ *in such a way that* $\gamma n^{1-2\kappa-\tau} \to \infty$ *as* $n \to \infty$*, we have, for some* $C > 0$*,*

$$\Pr(\mathcal{M}_o \subset \mathcal{M}_\gamma) = 1 - O\left(\exp[-C \min\{ \min_{1 \leqslant i \leqslant K} s_i, \min_{1 \leqslant i \leqslant K} n_{o(i)}^{1-2\kappa} / \log(n_{o(i)}) \}] \right),$$

where $\mathcal{M}_\gamma = \{1 \leqslant i \leqslant p : \ |\omega_i| \text{ is among the first } [\gamma p] \text{ largest of all}\}$.

Proof. The lemma can be proved similarly to the argument of Fan and Lv (2008). The key difference in proof is that the Lemma 4 and Lemma 5 in Fan and Lv (2008) have different conclusions in our proof. In our proof, Lemma 4 is that for $i = 1, \ldots, K$ and any $C > 0$, there are constants c_1 and c_2 with $0 < c_1 < 1 < c_2$ such that

$$\Pr\left(\langle \mathbf{S}_{(i)}\mathbf{e}_1, \mathbf{e}_1\rangle < c_1 n_{o(i)}/s_i \ \text{ or } \ > c_2 n_{o(i)}/s_i\right) \leqslant 4\exp\{-C\min(n_{o(i)}, s_i)\},$$

where $\mathbf{S}_{(i)} = (\widetilde{\mathbf{X}}_{o(i)}^\top \widetilde{\mathbf{X}}_{o(i)})^+ \widetilde{\mathbf{X}}_{o(i)}^\top \widetilde{\mathbf{X}}_{o(i)}$ and $\mathbf{e}_1 \in \mathbb{R}^{p_i}$ is a unit vector with the 1th entry 1 and 0 elsewhere. Lemma 5 is that let $\mathbf{S}_{(i)}\mathbf{e}_1 = (V_{(i)1}, \ldots, V_{(i)s_i})^\top$ for $i = 1, \ldots, K$, given that the first co-ordinate $V_{(i)1} = v$, the random vector $(V_{(i)2}, \ldots, V_{(i)s_i})^\top$ is uniformly distributed on the sphere $S^{s_i-2}(v - v^2)^{-1/2}$; moreover, for any $C > 0$, there is some $c > 1$ such that

$$\Pr\left(|V_{(i)2}| > c n_{o(i)}^{1/2} s_i^{-1} |A|\right) \leqslant 3\exp\{-C\min(n_{o(i)}, s_i - 1)\},$$

where A is an independent $N(0,1)$-distributed random variable. With the Lemma 4 and Lemma 5, we are able to prove this lemma similarly to Fan and Lv (2008).

S3 Proof of Theorems

For the proof of theorems, we introduce some notations as follows. Let $\|\mathbf{A}\| = \{\operatorname{tr}(\mathbf{A}^\top \mathbf{A})\}^{1/2}$ denote the norm of matrix $\mathbf{A}$, T_j be the collection of row indices in the jth data group $\mathbf{Z}_{(j)}$, and M_{oj} and M_{mj} be the collections of column indices of observed data and missing data in the jth data group $\mathbf{Z}_{(j)}$, respectively. Denote $\delta = \min(\sqrt{n_1}, \sqrt{q})$, $\mathbf{H}_1 = (\mathbf{\Lambda}^\top \mathbf{\Lambda}/q)(\mathbf{W}_{(1)}^\top \widetilde{\mathbf{W}}_{(1)}/n_1)\widetilde{\mathbf{V}}_{(1)}^{-1}$, and $\widetilde{\mathbf{V}}_{(1)}$ as the $r \times r$ diagonal matrix of the first r largest eigenvalues of $\mathbf{Z}_{(1)}\mathbf{Z}_{(1)}^\top/(n_1 q)$ in decreasing order. Let $\widehat{\mathbf{D}}_t = \{(d_{ti} - e_{ti})\delta_{ti}; i = 1, \ldots, q\}$ for $t = 1, \ldots, n$, where $\delta_{ti} = 1$ if Z_{ti} is missing and $\delta_{ti} = 0$ otherwise, and $d_{ti} = \widetilde{\boldsymbol{\lambda}}_i^\top \widetilde{\mathbf{w}}_t - \boldsymbol{\lambda}_i^\top \mathbf{w}_t$.

We provide the proof of some lemmas and theorems. Lemma 2, 4 and 5 are the lemmas from Bai (2003), and Lemma 3 is theorem 1 in Bai and Ng (2002), which are needed subsequently in the proof of theorems.

Lemma 2. *Under Assumptions (C1)-(C4), as $n_1, q \to \infty$:*

(i) $n_1^{-1}\widetilde{\mathbf{W}}_{(1)}^\top\{\mathbf{Z}_{(1)}\mathbf{Z}_{(1)}^\top/(n_1 q)\}\widetilde{\mathbf{W}}_{(1)} = \widetilde{\mathbf{V}}_{(1)} \xrightarrow{\text{p}} \mathbf{V}$,

(ii) $\frac{\widetilde{\mathbf{W}}_{(1)}^\top \mathbf{W}_{(1)}}{n_1}(\frac{\mathbf{\Lambda}^\top\mathbf{\Lambda}}{q})\frac{\mathbf{W}_{(1)}^\top \widetilde{\mathbf{W}}_{(1)}}{n_1} \xrightarrow{\text{p}} \mathbf{V}$,

where $\mathbf{V}$ is the diagonal matrix consisting of the eigenvalues of $\mathbf{\Sigma}_\Lambda \mathbf{\Sigma}_{\text{w}}$.

Lemma 3. *Under Assumption (C1)-(C4), we have*

$$n_1^{-1}\sum_{u=1}^{n_1}\|\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^{\top}\mathbf{W}_{(1)u}\|^2=O_p(\delta^{-2}).$$

Lemma 4. *Under Assumption (C1)-(C5), we have*

$$n_1^{-1}(\widetilde{\mathbf{W}}_{(1)}-\mathbf{W}_{(1)}\mathbf{H}_1)^{\top}\mathbf{W}_{(1)}=O_p(\delta^{-2}).$$

Lemma 5. *Under Assumption (C1)-(C5), we have*

$$n_1^{-1}\sum_{u=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^{\top}\mathbf{W}_{(1)u})\boldsymbol{e}_{(1)ui}=O_p(\delta^{-2})\ \text{ for }\ i=1,\ldots,q.$$

Lemma 6. *Under Assumption (C1)-(C5), we have*

$$\sum_{i\in M_{oj}}(\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)e_{li}=O_p\left(\frac{q_j}{\min(\sqrt{n_1},q)}\right)\ \text{ for }\ l\in T_j, j=2,\ldots,k.$$

Proof. From $\widetilde{\boldsymbol{\Lambda}}=\mathbf{Z}_{(1)}^{\top}\widetilde{\mathbf{W}}_{(1)}/n_1$ and $\mathbf{Z}_{(1)}=\mathbf{W}_{(1)}\boldsymbol{\Lambda}^{\top}+\boldsymbol{e}_{(1)}$, we have $\widetilde{\boldsymbol{\Lambda}}=\boldsymbol{\Lambda}\mathbf{W}_{(1)}^{\top}\widetilde{\mathbf{W}}_{(1)}/n_1+\boldsymbol{e}_{(1)}^{\top}\widetilde{\mathbf{W}}_{(1)}/n_1$. Writing $\mathbf{W}_{(1)}=\mathbf{W}_{(1)}-\widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1}+\widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1}$ and $\widetilde{\mathbf{W}}_{(1)}^{\top}\widetilde{\mathbf{W}}_{(1)}/n_1=\mathbf{I}_r$, we obtain

$$\begin{aligned}
\widetilde{\boldsymbol{\lambda}}_i &= \widetilde{\mathbf{W}}_{(1)}^{\top}\mathbf{W}_{(1)}\boldsymbol{\lambda}_i/n_1+\textstyle\sum_{t=1}^{n_1}\widetilde{\mathbf{W}}_{(1)t}\boldsymbol{e}_{(1)ti}/n_1\\
&= \widetilde{\mathbf{W}}_{(1)}^{\top}(\mathbf{W}_{(1)}-\widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1})\boldsymbol{\lambda}_i/n_1+\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\\
&\quad+\textstyle\sum_{t=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)t}-\mathbf{H}_1^{\top}\mathbf{W}_{(1)t})\boldsymbol{e}_{(1)ti}/n_1+\mathbf{H}_1^{\top}\sum_{t=1}^{n_1}\mathbf{W}_{(1)t}\boldsymbol{e}_{(1)ti}/n_1\\
&= (\widetilde{\mathbf{W}}_{(1)}-\mathbf{W}_{(1)}\mathbf{H}_1)^{\top}(\mathbf{W}_{(1)}-\widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1})\boldsymbol{\lambda}_i/n_1\\
&\quad+\mathbf{H}_1^{\top}\mathbf{W}_{(1)}^{\top}(\mathbf{W}_{(1)}-\widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1})\boldsymbol{\lambda}_i/n_1+\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\\
&\quad+\textstyle\sum_{t=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)t}-\mathbf{H}_1^{\top}\mathbf{W}_{(1)t})\boldsymbol{e}_{(1)ti}/n_1+\mathbf{H}_1^{\top}\sum_{t=1}^{n_1}\mathbf{W}_{(1)t}\boldsymbol{e}_{(1)ti}/n_1.
\end{aligned}$$

Thus, we have

$$
\begin{aligned}
\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i &= \mathbf{H}_1^\top \textstyle\sum_{u=1}^{n_1} \mathbf{W}_{(1)u}\boldsymbol{e}_{(1)ui}/n_1 \\
&\quad +(\widetilde{\mathbf{W}}_{(1)} - \mathbf{W}_{(1)}\mathbf{H}_1)^\top(\mathbf{W}_{(1)} - \widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1})\boldsymbol{\lambda}_i/n_1 \\
&\quad +\mathbf{H}_1^\top\mathbf{W}_{(1)}^\top(\mathbf{W}_{(1)} - \widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1})\boldsymbol{\lambda}_i/n_1 \\
&\quad +\textstyle\sum_{u=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)u} - \mathbf{H}_1^\top\mathbf{W}_{(1)u})\boldsymbol{e}_{(1)ui}/n_1.
\end{aligned} \tag{S3.1}
$$

Thus, $\sum_{i\in M_{oj}}(\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)e_{li} = I_1 + I_2 + I_3 + I_4$, where

$$
\begin{aligned}
I_1 &= \textstyle\sum_{i\in M_{oj}} \mathbf{H}_1^\top \sum_{u=1}^{n_1} \mathbf{W}_{(1)u}\boldsymbol{e}_{(1)ui}e_{li}/n_1, \\
I_2 &= \textstyle\sum_{i\in M_{oj}}(\widetilde{\mathbf{W}}_{(1)} - \mathbf{W}_{(1)}\mathbf{H}_1)^\top(\mathbf{W}_{(1)} - \widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1})\boldsymbol{\lambda}_i e_{li}/n_1, \\
I_3 &= \textstyle\sum_{i\in M_{oj}} \mathbf{H}_1^\top\mathbf{W}_{(1)}^\top(\mathbf{W}_{(1)} - \widetilde{\mathbf{W}}_{(1)}\mathbf{H}_1^{-1})\boldsymbol{\lambda}_i e_{li}/n_1, \\
I_4 &= \textstyle\sum_{i\in M_{oj}}\sum_{u=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)u} - \mathbf{H}_1^\top\mathbf{W}_{(1)u})\boldsymbol{e}_{(1)ui}e_{li}/n_1.
\end{aligned}
$$

Note that $\|\mathbf{H}_1\| = O_p(1)$ because

$$
\|\mathbf{H}_1\| \leqslant \|\frac{\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}}{q}\| \left\|\frac{\mathbf{W}_{(1)}^\top\mathbf{W}_{(1)}}{n_1}\right\|^{1/2} \left\|\frac{\widetilde{\mathbf{W}}_{(1)}^\top\widetilde{\mathbf{W}}_{(1)}}{n_1}\right\|^{1/2} \|\widetilde{\mathbf{V}}_{(1)}^{-1}\|
$$

and each of the matrix norms is stochastically bounded by Assumption (C1) and (C2) together with $\widetilde{\mathbf{W}}_{(1)}^\top\widetilde{\mathbf{W}}_{(1)}/n_1 = \mathbf{I}_r$ and Lemma 2. By Assumption (C3) and (C4), we have

$$
\begin{aligned}
\|I_1\| &\leqslant \|\mathbf{H}_1^\top\|\tfrac{q_j}{\sqrt{n_1}}\left\{\tfrac{1}{q_j}\textstyle\sum_{i\in M_{oj}}\|\tfrac{1}{\sqrt{n_1}}\sum_{u=1}^{n_1}\mathbf{W}_{(1)u}\boldsymbol{e}_{(1)ui}\|^2\right\}^{1/2}\left\{\tfrac{1}{q_j}\textstyle\sum_{i\in M_{oj}} e_{li}^2\right\}^{1/2} \\
&= O_p\left(\tfrac{q_j}{\sqrt{n_1}}\right).
\end{aligned}
$$

By Assumption (C2) and (C3), we have $\sum_{i\in M_{oj}}\boldsymbol{\lambda}_i e_{li}/\sqrt{q_j} = O_p(1)$ since $E(\sum_{i\in M_{oj}}\boldsymbol{\lambda}_i e_{li}/\sqrt{q_j}) = 0$ and $E(\|\sum_{i\in M_{oj}}\boldsymbol{\lambda}_i e_{li}/\sqrt{q_j}\|^2) \leqslant \bar{\lambda}^2\sum_{i\in M_{oj}}\tau_{ii}/q_j \leqslant$

C. By Lemma 3, we obtain

$$\begin{aligned}
\|I_2\| &= \|\tfrac{1}{n_1}\textstyle\sum_{u=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u})(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u})^\top\| \\
&\quad\times\|\mathbf{H}_1^{-1}\|\|\textstyle\sum_{i\in M_{oj}}\boldsymbol{\lambda}_i e_{li}\| \\
&\leqslant (\tfrac{1}{n_1}\textstyle\sum_{u=1}^{n_1}\|\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u}\|^2)\|\mathbf{H}_1^{-1}\|\sqrt{q_j}\|\tfrac{1}{\sqrt{q_j}}\textstyle\sum_{i\in M_{oj}}\boldsymbol{\lambda}_i e_{li}\| \\
&= O_p(\tfrac{\sqrt{q_j}}{\delta^2}).
\end{aligned}$$

By Lemma 4, we have

$$\begin{aligned}
\|I_3\| &\leqslant \|\mathbf{H}_1^\top\|\|\tfrac{1}{n_1}\textstyle\sum_{u=1}^{n_1}\mathbf{W}_{(1)u}(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^{-1}\mathbf{W}_{(1)u})^\top\|\|\mathbf{H}_1^{-1}\| \\
&\quad\times\sqrt{q_j}\|\tfrac{1}{\sqrt{q_j}}\textstyle\sum_{i\in M_{oj}}\boldsymbol{\lambda}_i e_{li}\| \\
&\leqslant O_p(\tfrac{\sqrt{q_j}}{\delta^2}).
\end{aligned}$$

By Lemma 5 and Assumption (C3), we have

$$\begin{aligned}
\|I_4\| &\leqslant \sqrt{q_j}\left\{\tfrac{1}{q_j}\textstyle\sum_{i\in M_{oj}}\|\tfrac{1}{n_1}\sum_{u=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u})\boldsymbol{e}_{(1)ui}\|^2\right\}^{1/2} \\
&\quad\times\sqrt{q_j}\left(\tfrac{1}{q_j}\textstyle\sum_{i\in M_{oj}}e_{li}^2\right)^{1/2} \\
&\leqslant O_p(\tfrac{q_j}{\delta^2}).
\end{aligned}$$

Thus, we obtain

$$\sum_{i\in M_{oj}}(\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)e_{li}=O_p\left(\frac{q_j}{\min(\sqrt{n_1},q)}\right).$$

Lemma 7. *Under Assumption (C1)-(C5), we have*

$$\sum_{i\in M_{oj}}\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2=O_p\left(\frac{q_j}{\min(n_1,q^2)}\right)\ \text{ for }\ j=2,\ldots,k.$$

Proof. Since equation (S3.1) and $(x+y+z+u)^2\leqslant 4(x^2+y^2+z^2+u^2)$, we have $\sum_{i\in M_{oj}}\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\leqslant 4(J_1+J_2+J_3+J_4)$, where

$$\begin{aligned}
J_1 &= \textstyle\sum_{i\in M_{oj}}\|\mathbf{H}_1^\top\sum_{u=1}^{n_1}\mathbf{W}_{(1)u}\boldsymbol{e}_{(1)ui}/n_1\|^2,\\
J_2 &= \textstyle\sum_{i\in M_{oj}}\|(\widetilde{\mathbf{W}}_{(1)}-\mathbf{W}_{(1)}\mathbf{H}_1)^\top(\widetilde{\mathbf{W}}_{(1)}-\mathbf{W}_{(1)}\mathbf{H}_1)\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i/n_1\|^2,\\
J_3 &= \textstyle\sum_{i\in M_{oj}}\|\mathbf{H}_1^\top\mathbf{W}_{(1)}^\top(\widetilde{\mathbf{W}}_{(1)}-\mathbf{W}_{(1)}\mathbf{H}_1)\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i/n_1\|^2,\\
J_4 &= \textstyle\sum_{i\in M_{oj}}\|\sum_{u=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u})\boldsymbol{e}_{(1)ui}/n_1\|^2.
\end{aligned}$$

By Assumption (C4), we have

$$J_1\leqslant\frac{1}{n_1}\|\mathbf{H}_1^\top\|^2q_j(\frac{1}{q_j}\sum_{i\in M_{oj}}\|\frac{1}{\sqrt{n_1}}\sum_{u=1}^{n_1}\mathbf{W}_{(1)u}\boldsymbol{e}_{(1)ui}\|^2)\leqslant O_p(\frac{q_j}{n_1}).$$

By Assumption (C2) and Lemma 3, we have

$$\begin{aligned}
J_2 &\leqslant \textstyle\|\frac{1}{n_1}\sum_{u=1}^{n_1}(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u})(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u})^\top\|^2\\
&\quad \textstyle\times\|\mathbf{H}_1^{-1}\|^2(\sum_{i\in M_{oj}}\|\boldsymbol{\lambda}_i\|^2)\\
&\leqslant \textstyle(\frac{1}{n_1}\sum_{u=1}^{n_1}\|\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u}\|^2)^2O_p(1)q_j\bar{\lambda}^2\\
&= O_p(q_j\delta^{-4}).
\end{aligned}$$

By Assumption (C2) and Lemma 4, we have

$$\begin{aligned}
J_3 &\leqslant \textstyle\|\mathbf{H}_1^\top\|^2\|\frac{1}{n_1}\sum_{u=1}^{n_1}\mathbf{W}_{(1)u}(\widetilde{\mathbf{W}}_{(1)u}-\mathbf{H}_1^\top\mathbf{W}_{(1)u})^\top\|^2\|\mathbf{H}_1^{-1}\|^2(\sum_{i\in M_{oj}}\|\boldsymbol{\lambda}_i\|^2)\\
&= O_p(q_j\delta^{-4}).
\end{aligned}$$

By Lemma 5, we have $J_4=O_p(q_j\delta^{-4})$. Then we obtain

$$\sum_{i\in M_{oj}}\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2=O_p\left(\frac{q_j}{\min(n_1,q^2)}\right).$$

Lemma 8. *Under Assumption (C1)-(C5), as $q,n_1\to\infty$, we have* *(i)* $\widetilde{\mathbf{W}}_{(1)t} - \mathbf{H}_1^\top \mathbf{W}_{(1)t} = O_p\left((\min(n_1, \sqrt{q}))^{-1}\right)$ *for* $t = 1, \ldots, n_1$,

(ii) $\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i = O_p\left((\min(\sqrt{n_1}, q))^{-1}\right)$ *for* $i = 1, \ldots, q$,

(iii) $\widetilde{\mathbf{W}}_{(j)t} - \mathbf{H}_1^\top \mathbf{W}_{(j)t} = O_p\left(\frac{q_j}{q\min(\sqrt{n_1}, \sqrt{q_j})}\right)$ *for* $j = 2, \ldots, k$ *and* $t = 1, \ldots, n_j$.

Proof. Since $\widetilde{\mathbf{W}}_{(1)}$ and $\widetilde{\boldsymbol{\lambda}}_i$ are obtained in the complete observed data block $\mathbf{Z}_{(1)}$, the proofs of part (i) and (ii) are similar to Bai (2003). The details are omitted.

Consider part (iii). Since $\widetilde{\mathbf{W}}_{(j)} = \mathbf{Z}_{o(j)}\widetilde{\boldsymbol{\Lambda}}_{o(j)}(\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}$ and $\mathbf{Z}_{o(j)} = \mathbf{W}_{(j)}\boldsymbol{\Lambda}_{o(j)}^\top + \boldsymbol{e}_{o(j)}$ for $j = 2, \ldots, k$, then we have, for $t = 1, \ldots, n_j$,

$$
\begin{aligned}
\widetilde{\mathbf{W}}_{(j)t} &= (\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \boldsymbol{\Lambda}_{o(j)}\mathbf{W}_{(j)t} + (\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \boldsymbol{e}_{o(j)t} \\
&= \mathbf{H}_1^\top \mathbf{W}_{(j)t} + (\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top (\boldsymbol{\Lambda}_{o(j)} - \widetilde{\boldsymbol{\Lambda}}_{o(j)}\mathbf{H}_1^\top)\mathbf{W}_{(j)t} \\
&\quad + (\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}(\widetilde{\boldsymbol{\Lambda}}_{o(j)} - \boldsymbol{\Lambda}_{o(j)}\mathbf{H}_1^{-1\top})^\top \boldsymbol{e}_{o(j)t} \\
&\quad + (\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}\mathbf{H}_1^{-1}\boldsymbol{\Lambda}_{o(j)}^\top \boldsymbol{e}_{o(j)t}.
\end{aligned}
$$

That is

$$
\begin{aligned}
&\widetilde{\mathbf{W}}_{(j)t} - \mathbf{H}_1^\top \mathbf{W}_{(j)t} \\
= &(\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}\{(\widetilde{\boldsymbol{\Lambda}}_{o(j)} - \boldsymbol{\Lambda}_{o(j)}\mathbf{H}_1^{-1\top})^\top (\boldsymbol{\Lambda}_{o(j)} - \widetilde{\boldsymbol{\Lambda}}_{o(j)}\mathbf{H}_1^\top)\mathbf{W}_{(j)t} \\
&+ \mathbf{H}_1^{-1}\boldsymbol{\Lambda}_{o(j)}^\top (\boldsymbol{\Lambda}_{o(j)} - \widetilde{\boldsymbol{\Lambda}}_{o(j)}\mathbf{H}_1^\top)\mathbf{W}_{(j)t} \\
&+ (\widetilde{\boldsymbol{\Lambda}}_{o(j)} - \boldsymbol{\Lambda}_{o(j)}\mathbf{H}_1^{-1\top})^\top \boldsymbol{e}_{o(j)t} + \mathbf{H}_1^{-1}\boldsymbol{\Lambda}_{o(j)}^\top \boldsymbol{e}_{o(j)t}\} \\
= &(\widetilde{\boldsymbol{\Lambda}}_{o(j)}^\top \widetilde{\boldsymbol{\Lambda}}_{o(j)})^{-1}(I_1 + I_2 + I_3 + I_4).
\end{aligned}
$$

By Lemma 2, we have

$$\begin{aligned}
\widetilde{\mathbf{\Lambda}}_{o(j)}^\top \widetilde{\mathbf{\Lambda}}_{o(j)} &= \textstyle\sum_{i=1}^{q_j} \widetilde{\boldsymbol{\lambda}}_{o(j)i} \widetilde{\boldsymbol{\lambda}}_{o(j)i}^\top \\
&\leqslant \textstyle\sum_{i=1}^{q} \widetilde{\boldsymbol{\lambda}}_i \widetilde{\boldsymbol{\lambda}}_i^\top = \widetilde{\mathbf{\Lambda}}^\top \widetilde{\mathbf{\Lambda}} = q\{\frac{1}{n_1} \widetilde{\mathbf{W}}_{(1)}^\top \frac{\mathbf{Z}_{(1)} \mathbf{Z}_{(1)}^\top}{n_1 q} \widetilde{\mathbf{W}}_{(1)}\} \\
&= O_p(q).
\end{aligned}$$

Thus, by Assumption (C1) and Lemma 7, we have

$$\begin{aligned}
\|I_1\| &= \| \textstyle\sum_{i \in M_{oj}} (\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1} \boldsymbol{\lambda}_i)(\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1} \boldsymbol{\lambda}_i)^\top \| \|\mathbf{H}_1^{-1}\| \|\mathbf{W}_{(j)t}\| \\
&\leqslant \textstyle\sum_{i \in M_{oj}} \|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1} \boldsymbol{\lambda}_i\|^2 O_p(1) \\
&= O_p(\frac{q_j}{\min(n_1, q^2)}).
\end{aligned}$$

By Assumption (C1)-(C2) and Lemma 7, we have

$$\begin{aligned}
\|I_2\| &= \|\mathbf{H}_1^{-1}\| \| \textstyle\sum_{i \in M_{oj}} \boldsymbol{\lambda}_i (\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1} \boldsymbol{\lambda}_i)^\top \| \|\mathbf{H}_1^\top\| \|\mathbf{W}_{(j)t}\| \\
&\leqslant O_p(1) (\textstyle\sum_{i \in M_{oj}} \|\boldsymbol{\lambda}_i\|^2)^{1/2} (\textstyle\sum_{i \in M_{oj}} \|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1} \widetilde{\boldsymbol{\lambda}}_i\|^2)^{1/2} \\
&\leqslant O_p(1) \sqrt{q_j} \bar{\lambda} O_p(\frac{\sqrt{q_j}}{\min(\sqrt{n_1}, q)}) \\
&= O_p(\frac{q_j}{\min(\sqrt{n_1}, q)}).
\end{aligned}$$

By Lemma 6, we have

$$\|I_3\| = \| \sum_{i \in M_{oj}} (\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1} \boldsymbol{\lambda}_i) e_{li} \| = O_p(\frac{q_j}{\min(\sqrt{n_1}, q)}) \ \text{ for } \ l \in T_j.$$

By Assumption (C2) and (C3), we have

$$\|I_4\| = \|\mathbf{H}_1^{-1}\| \sqrt{q_j} \| \frac{1}{\sqrt{q_j}} \sum_{i \in M_{oj}} \boldsymbol{\lambda}_i e_{li} \| \leqslant O_p(\sqrt{q_j}).$$

Thus, we have

$$\widetilde{\mathbf{W}}_{(j)t} - \mathbf{H}_1^\top \mathbf{W}_{(j)t} \leqslant O_p(\frac{q_j}{q \min(\sqrt{n_1}, \sqrt{q_j})}).$$

The proof of part (iii) is completed.

Proof of Theorem 1. Based on Lemma 8, we can obtain $\widetilde{\boldsymbol{\lambda}}_i^\top \widetilde{\mathbf{w}}_t - \boldsymbol{\lambda}_i^\top \mathbf{w}_t = (\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)^\top(\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t) + (\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)^\top \mathbf{H}_1^\top \mathbf{w}_t + \boldsymbol{\lambda}_i^\top \mathbf{H}_1^{-1\top}(\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t) = O_p\{\sqrt{q_j}(\min(\sqrt{n_1 q_j}, q))^{-1}\}$. The proof of theorem is completed.

Lemma 9. *Under Assumption (C1), (C2) and (C3), we have*

$$\|\boldsymbol{\Lambda}^\top \widehat{\mathbf{D}}_t/q\| = O_p\left((\min(\sqrt{n_1}, \sqrt{q}))^{-1}\right) \qquad \text{for } \ t \in T_j, j = 2, \ldots, k.$$

Proof. Based on the definition of $\widehat{\mathbf{D}}_t$, we have for $t \in T_j$ $(j = 2, \ldots, k)$

$$\begin{aligned} \|\boldsymbol{\Lambda}^\top \widehat{\mathbf{D}}_t/q\| &= \|\textstyle\sum_{i=1}^q \boldsymbol{\lambda}_i(d_{ti} - e_{ti})\delta_{ti}/q\| = \|\sum_{i\in M_{mj}} \boldsymbol{\lambda}_i(d_{ti} - e_{ti})/q\| \\ &\leqslant \|\textstyle\sum_{i\in M_{mj}} \boldsymbol{\lambda}_i d_{ti}/q\| + \|\sum_{i\in M_{mj}} \boldsymbol{\lambda}_i e_{ti}/q\|. \end{aligned}$$

Based on Assumption (C1) and (C2), we have

$$\begin{aligned} \|\tfrac{1}{q}\textstyle\sum_{i\in M_{mj}} \boldsymbol{\lambda}_i d_{ti}\| &= \|\tfrac{1}{q}\textstyle\sum_{i\in M_{mj}} \boldsymbol{\lambda}_i(\widetilde{\boldsymbol{\lambda}}_i^\top \widetilde{\mathbf{w}}_t - \boldsymbol{\lambda}_i^\top \mathbf{w}_t)\| \\ &= \|\tfrac{1}{q}\textstyle\sum_{i\in M_{mj}} \boldsymbol{\lambda}_i\{(\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)^\top(\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t) \\ &\quad +(\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)^\top \mathbf{H}_1^\top \mathbf{w}_t \\ &\quad +\boldsymbol{\lambda}_i^\top \mathbf{H}_1^{-1\top}(\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t)\}\| \\ &\leqslant \bar{\lambda}\tfrac{q-q_j}{q}\tfrac{1}{q-q_j}\textstyle\sum_{i\in M_{mj}} \|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|\|\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t\| \\ &\quad +\bar{\lambda}\tfrac{q-q_j}{q}\tfrac{1}{q-q_j}\textstyle\sum_{i\in M_{mj}} \|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|\|\mathbf{H}_1\|\|\mathbf{w}_t\| \\ &\quad +\|\tfrac{1}{q}\textstyle\sum_{i=1}^q \boldsymbol{\lambda}_i\boldsymbol{\lambda}_i^\top\|\|\mathbf{H}_1^{-1}\|\|\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t\| \\ &\leqslant O_p(\tfrac{1}{\sqrt{n_1}}) + O_p(\tfrac{\sqrt{q_j}}{q}). \end{aligned}$$

For Assumption (C2) and (C3), we have $\|\frac{1}{q}\sum_{i\in M_{mj}} \boldsymbol{\lambda}_i e_{ti}\| = O_p(\frac{1}{\sqrt{q}})$. Thus, we obtain $\|\frac{1}{q}\mathbf{\Lambda}^\top\widehat{\mathbf{D}}_t\| = O_p(\frac{1}{\min(\sqrt{n_1},\sqrt{q})})$. The proof of the lemma is completed.

Lemma 10. *Under Assumption (C1)-(C5), we have*

$$\frac{1}{nq^2}\sum_{l=1}^{n}\|\boldsymbol{e}_l^\top\widehat{\mathbf{D}}_t\|^2 = O_p\left((\min(n_1,q))^{-1}\right) \qquad \text{for } \ t\in T_j, j=2,\ldots,k.$$

Proof. Based on the definition of $\widehat{\mathbf{D}}_t$, we have for $t\in T_j, j=2,\ldots,k$

$$\begin{aligned}\tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\boldsymbol{e}_l^\top\widehat{\mathbf{D}}_t\|^2 &= \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}e_{li}(d_{ti}-e_{ti})\delta_{ti}\|^2\\ &= \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i\in M_{mj}}e_{li}(d_{ti}-e_{ti})\|^2\\ &\leqslant \tfrac{2}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i\in M_{mj}}e_{li}d_{ti}\|^2\\ &\quad +\tfrac{2}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i\in M_{mj}}e_{li}e_{ti}\|^2.\end{aligned} \tag{S3.2}$$

For Assumption (C3), we have

$$\frac{1}{nq^2}\sum_{l=1}^{n}\|\sum_{i\in M_{mj}}e_{li}d_{ti}\|^2 \leqslant \left\{\frac{1}{nq}\sum_{l=1}^{n}\sum_{i\in M_{mj}}e_{li}^2\right\}\frac{1}{q}\sum_{i\in M_{mj}}d_{ti}^2 = O_p(1)\frac{1}{q}\sum_{i\in M_{mj}}d_{ti}^2.$$

Based on Lemma 7 and 8, we can obtain

$$\begin{aligned}\tfrac{1}{q}\textstyle\sum_{i\in M_{mj}}d_{ti}^2 &= \tfrac{1}{q}\textstyle\sum_{i\in M_{mj}}(\widetilde{\boldsymbol{\lambda}}_i^\top\widetilde{\mathbf{w}}_t-\boldsymbol{\lambda}_i^\top\mathbf{w}_t)^2\\ &= \tfrac{1}{q}\textstyle\sum_{i\in M_{mj}}\{(\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)^\top(\widetilde{\mathbf{w}}_t-\mathbf{H}_1^\top\mathbf{w}_t)\\ &\quad +(\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i)^\top\mathbf{H}_1^\top\mathbf{w}_t+\boldsymbol{\lambda}_i^\top\mathbf{H}_1^{-1\top}(\widetilde{\mathbf{w}}_t-\mathbf{H}_1^\top\mathbf{w}_t)\}^2\\ &\leqslant \tfrac{C(q-q_j)}{q}\{(\tfrac{1}{q-q_j}\textstyle\sum_{i\in M_{mj}}\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2)\|\widetilde{\mathbf{w}}_t-\mathbf{H}_1^\top\mathbf{w}_t\|^2\\ &\quad +(\tfrac{1}{q-q_j}\textstyle\sum_{i\in M_{mj}}\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2)\|\mathbf{H}_1\|^2\|\mathbf{w}_t\|^2\\ &\quad +\bar{\lambda}^2\|\mathbf{H}_1^{-1}\|^2\|\widetilde{\mathbf{w}}_t-\mathbf{H}_1^\top\mathbf{w}_t\|^2\}\\ &\leqslant O_p(\tfrac{1}{n_1})+O_p(\tfrac{q_j}{q^2}).\end{aligned} \tag{S3.3}$$

Thus, we have $\frac{1}{nq^2}\sum_{l=1}^{n}\|\sum_{i\in M_{mj}} e_{li}d_{ti}\|^2 \leqslant O_p(\frac{1}{n_1})+O_p(\frac{q_j}{q^2})$. For Assumption (C3), we have

$$\begin{aligned} E\{\tfrac{1}{nq}\textstyle\sum_{l=1}^{n}\|\sum_{i\in M_{mj}} e_{li}e_{ti}\|^2\} &= \tfrac{1}{nq}\textstyle\sum_{l=1}^{n}\sum_{i\in M_{mj}}\sum_{u\in M_{mj}} E(e_{li}e_{ti}e_{lu}e_{tu}) \\ &= \tfrac{1}{q}\textstyle\sum_{i\in M_{mj}}\sum_{u\in M_{mj}}\tau_{iu}^2 \\ &\quad +\tfrac{1}{nq}\textstyle\sum_{i\in M_{mj}}\sum_{u\in M_{mj}}\{E(e_{ti}^2e_{tu}^2)-\tau_{iu}^2\} \\ &\leqslant C. \end{aligned}$$

Thus, $\frac{1}{nq^2}\sum_{l=1}^{n}\|\sum_{i\in M_{mj}} e_{li}e_{ti}\|^2 = O_p(\frac{1}{q})$. From equation (S3.2), we have $\frac{1}{nq^2}\sum_{l=1}^{n}\|\boldsymbol{e}_l^\top\widehat{\mathbf{D}}_t\|^2 = O_p(\frac{1}{\min(n_1,q)})$. The proof of lemma is completed.

Lemma 11. *Under Assumptions (C1)-(C5), we have*

$$\frac{1}{nq}\sum_{l=1}^{n}\|\boldsymbol{\Lambda}^\top\widehat{\mathbf{D}}_l\|^2 = O_p\left(\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n}\right).$$

Proof. Based on the definition of $\widehat{\mathbf{D}}_l$, we have

$$\frac{1}{nq}\sum_{l=1}^{n}\|\boldsymbol{\Lambda}^\top\widehat{\mathbf{D}}_l\|^2 \leqslant \frac{2}{nq}\sum_{l=1}^{n}\|\sum_{i=1}^{q}\boldsymbol{\lambda}_i d_{li}\delta_{li}\|^2+\frac{2}{nq}\sum_{l=1}^{n}\|\sum_{i=1}^{q}\boldsymbol{\lambda}_i e_{li}\delta_{li}\|^2. \quad \text{(S3.4)}$$

Then, we have

$$\begin{aligned} &\tfrac{1}{nq}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}\boldsymbol{\lambda}_i d_{li}\delta_{li}\|^2 \\ \leqslant\ &\tfrac{C}{nq}\textstyle\sum_{l=1}^{n}\sum_{i=1}^{q}\|\boldsymbol{\lambda}_i d_{li}\|^2\delta_{li} \\ \leqslant\ &\tfrac{C}{nq}\textstyle\sum_{l=1}^{n}\sum_{i=1}^{q}\|\boldsymbol{\lambda}_i\|^2\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\|\widetilde{\mathbf{w}}_l-\mathbf{H}_1^\top\mathbf{w}_l\|^2\delta_{li} \\ &+\tfrac{C}{nq}\textstyle\sum_{l=1}^{n}\sum_{i=1}^{q}\|\boldsymbol{\lambda}_i\|^2\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\|\mathbf{H}_1^\top\|^2\|\mathbf{w}_l\|^2\delta_{li} \\ &+\tfrac{C}{nq}\textstyle\sum_{l=1}^{n}\sum_{i=1}^{q}\|\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i^\top\|^2\|\mathbf{H}_1^{-1}\|^2\|\widetilde{\mathbf{w}}_l-\mathbf{H}_1^\top\mathbf{w}_l\|^2\delta_{li} \\ =\ &I_1+I_2+I_3. \end{aligned}$$

For Lemma 7 and 8, we have

$$\begin{aligned}
I_1 &\leqslant \tfrac{C\bar{\lambda}^2}{nq}\sum_{l=1}^n(\sum_{i=1}^q\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2)\|\widetilde{\mathbf{w}}_l-\mathbf{H}_1^\top\mathbf{w}_l\|^2\delta_{li}\\
&= \tfrac{C\bar{\lambda}^2}{nq}\sum_{l=1}^n\sum_{j=2}^k(\sum_{i\in M_{mj}}\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2)\|\widetilde{\mathbf{w}}_l-\mathbf{H}_1^\top\mathbf{w}_l\|^2\mathrm{I}(l\in T_j)\\
&\leqslant \tfrac{1}{nq}\sum_{l=1}^n\sum_{j=2}^k O_p(\tfrac{(q-q_j)q_j^2}{q^2\min(n_1,q_j)\min(n_1,q^2)})\mathrm{I}(l\in T_j)\\
&= \sum_{j=2}^k O_p(\tfrac{n_jq_j^2}{nq^2\min(n_1,q_j)\min(n_1,q^2)}),
\end{aligned}$$

where $\mathrm{I}(\cdot)$ is a indicator function. Similarly, we have

$$\begin{aligned}
I_2 &\leqslant \tfrac{C\bar{\lambda}^2}{nq}\sum_{l=1}^n\sum_{i=1}^q\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\delta_{li}\\
&= \tfrac{C\bar{\lambda}^2}{nq}\sum_{l=1}^n\sum_{j=2}^k\{\sum_{i\in M_{mj}}\|\widetilde{\boldsymbol{\lambda}}_i-\mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\mathrm{I}(l\in T_j)\}\\
&\leqslant \sum_{j=2}^k O_p(\tfrac{n_j}{n\min(n_1,q^2)}),
\end{aligned}$$

and

$$I_3\leqslant\frac{C}{nq}\sum_{i=1}^q\|\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i^\top\|^2\sum_{l=1}^n\sum_{j=2}^k\|\widetilde{\mathbf{w}}_l-\mathbf{H}_1^\top\mathbf{w}_l\|^2\mathrm{I}(l\in T_j)\leqslant\sum_{j=2}^k O_p(\frac{n_jq_j^2}{nq^2\min(n_1,q_j)}).$$

Thus, $\frac{1}{nq}\sum_{l=1}^n\|\sum_{i=1}^q\boldsymbol{\lambda}_i d_{li}\delta_{li}\|^2\leqslant O_p(\frac{\max_{2\leqslant j\leqslant k}(n_j)}{nn_1})+O_p(\frac{\max_{2\leqslant j\leqslant k}(n_jq_j)}{nq^2})\}$. For Assumptions (C2) and (C3), we have

$$\begin{aligned}
&E\{\tfrac{1}{nq}\sum_{l=1}^n\|\sum_{i=1}^q\boldsymbol{\lambda}_i e_{li}\delta_{li}\|^2\}\\
=\ &\tfrac{1}{nq}\sum_{l=1}^n\sum_{i=1}^q\sum_{u=1}^q E(\boldsymbol{\lambda}_i^\top\boldsymbol{\lambda}_u e_{li}e_{lu}\delta_{li}\delta_{lu})\\
\leqslant\ &\tfrac{\bar{\lambda}^2}{nq}\sum_{l=1}^n\sum_{i=1}^q\sum_{u=1}^q E(e_{li}e_{lu})\delta_{li}\delta_{lu}\\
=\ &\tfrac{\bar{\lambda}^2}{nq}\sum_{l=1}^n\sum_{j=2}^k\{\sum_{i\in M_{mj}}\sum_{u\in M_{mj}}E(e_{li}e_{lu})\}\mathrm{I}(l\in T_j)\\
=\ &\bar{\lambda}^2\sum_{j=2}^k\tfrac{n_j}{nq}\sum_{i\in M_{mj}}\sum_{u\in M_{mj}}|\tau_{iu}|\\
\leqslant\ &O_p(\tfrac{\max_{2\leqslant j\leqslant k}(n_j)}{n}).
\end{aligned}$$

From inequality (S3.4), we have $\frac{1}{nq}\sum_{l=1}^{n}\|\boldsymbol{\Lambda}^{\top}\widehat{\mathbf{D}}_l\|^2 = O_p(\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n})$. The proof of lemma is completed.

Lemma 12. *Under Assumptions (C1)-(C5), we have*

$$\frac{1}{nq^2}\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^{\top}\boldsymbol{e}_t\|^2 = O_p\left\{\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n\min(n_1,q)}\right\} \quad \text{for} \ \ t=1,\dots,n.$$

Proof. Based on the definition of $\widehat{\mathbf{D}}_l$, we have

$$\frac{1}{nq^2}\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^{\top}\boldsymbol{e}_t\|^2 \leqslant \frac{2}{nq^2}\sum_{l=1}^{n}\|\sum_{i=1}^{q}d_{li}e_{ti}\delta_{li}\|^2+\frac{2}{nq^2}\sum_{l=1}^{n}\|\sum_{i=1}^{q}e_{li}e_{ti}\delta_{li}\|^2, \quad \text{(S3.5)}$$

where $\frac{1}{nq^2}\sum_{l=1}^{n}\|\sum_{i=1}^{q}d_{li}e_{ti}\delta_{li}\|^2 \leqslant \frac{1}{nq}\sum_{l=1}^{n}(\sum_{i=1}^{q}d_{li}^2\delta_{li})(\frac{1}{q}\sum_{i=1}^{q}e_{ti}^2)$. Based on Assumption (C3), we know $E\{\frac{1}{q}\sum_{i=1}^{q}e_{ti}^2\} \leqslant C$. From (S3.3), we have

$$\begin{aligned}\frac{1}{nq}\sum_{l=1}^{n}\sum_{i=1}^{q}d_{li}^2\delta_{li} &= \frac{1}{n}\sum_{l=1}^{n}\sum_{j=2}^{k}(\frac{1}{q}\sum_{i\in M_{mj}}d_{li}^2)\mathrm{I}(l\in T_j)\\ &\leqslant O_p\{\frac{\max_{2\leqslant j\leqslant k}(n_j)}{nn_1}\}+O_p\{\frac{\max_{2\leqslant j\leqslant k}(n_jq_j)}{nq^2}\}.\end{aligned}$$

Thus, $\frac{1}{nq^2}\sum_{l=1}^{n}\|\sum_{i=1}^{q}d_{li}e_{ti}\delta_{li}\|^2 \leqslant O_p\{\frac{\max_{2\leqslant j\leqslant k}(n_j)}{nn_1}\} + O_p\{\frac{\max_{2\leqslant j\leqslant k}(n_jq_j)}{nq^2}\}$.

Based on Assumption (C3), we have

$$\begin{aligned}&E\{\frac{1}{nq}\sum_{l=1}^{n}\|\sum_{i=1}^{q}e_{li}e_{ti}\delta_{li}\|^2\}\\ =\ & \frac{1}{nq}\sum_{l=1}^{n}\sum_{i=1}^{q}\sum_{u=1}^{q}E(e_{li}e_{ti}e_{lu}e_{tu})\delta_{li}\delta_{lu}\\ =\ & \frac{1}{nq}\sum_{l=1}^{n}\sum_{j=2}^{k}\{\sum_{i\in M_{mj}}\sum_{u\in M_{mj}}E(e_{li}e_{ti}e_{lu}e_{tu})\mathrm{I}(l\in T_j)\}\\ \leqslant\ & \sum_{j=2}^{k}\{\frac{n_j}{nq}\sum_{i\in M_{mj}}\sum_{u\in M_{mj}}\tau_{iu}^2+\frac{1}{nq}\sum_{i\in M_{mj}}\sum_{u\in M_{mj}}(E(e_{ti}^2e_{tu}^2)-\tau_{iu}^2)\}\\ \leqslant\ & O_p(\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n}).\end{aligned}$$

Thus, $\frac{1}{nq^2}\sum_{l=1}^{n}\|\sum_{i=1}^{q}e_{li}e_{ti}\delta_{li}\|^2 = O_p(\frac{\max_{2\leqslant j\leqslant k}(n_j)}{nq})$. From inequality (S3.5), we have $\frac{1}{nq^2}\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^\top \boldsymbol{e}_t\|^2 = O_p(\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n\min(n_1,q)})$. The proof of lemma is completed.

Lemma 13. *Under Assumptions (C1)-(C5), we have*

$$\frac{1}{nq^2}\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^\top\widehat{\mathbf{D}}_t\|^2 = O_p\left\{\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n\min(n_1,q)}\right\} \text{ for } t\in T_u, u=2,\ldots,k.$$

Proof. Based on the definition of $\widehat{\mathbf{D}}_l$, we have

$$\begin{aligned}
\tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^\top\widehat{\mathbf{D}}_t\|^2 &= \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}\widehat{\mathbf{D}}_{li}\widehat{\mathbf{D}}_{ti}\|^2\\
&= \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}(d_{li}-e_{li})(d_{ti}-e_{ti})\delta_{li}\delta_{ti}\|^2\\
&\leqslant C\{\tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}d_{li}d_{ti}\delta_{li}\delta_{ti}\|^2\\
&\quad+\tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}d_{li}e_{ti}\delta_{li}\delta_{ti}\|^2\\
&\quad+\tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}e_{li}d_{ti}\delta_{li}\delta_{ti}\|^2\\
&\quad+\tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\|\sum_{i=1}^{q}e_{li}e_{ti}\delta_{li}\delta_{ti}\|^2\}\\
&= C(I_1+I_2+I_3+I_4).
\end{aligned}$$

From (S3.3), we have

$$\begin{aligned}
I_1 &\leqslant \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}(\sum_{i=1}^{q}d_{li}^2\delta_{li})(\sum_{i=1}^{q}d_{ti}^2\delta_{ti})\\
&= \tfrac{1}{n}\textstyle\sum_{l=1}^{n}\sum_{j=2}^{k}(\tfrac{1}{q}\sum_{i\in M_{mj}}d_{li}^2)\mathrm{I}(l\in T_j)(\tfrac{1}{q}\sum_{i\in M_{mu}}d_{ti}^2)\\
&\leqslant O_p(\tfrac{\max_{2\leqslant j\leqslant k}(n_j)}{nn_1^2})+O(\tfrac{\max_{2\leqslant j\leqslant k}(n_jq_j)}{nn_1q^2})+O_p(\tfrac{\max_{2\leqslant j\leqslant k}(n_jq_j)q_u}{nq^4}).
\end{aligned}$$

Based on Assumption (C3) and (S3.3), we have

$$\begin{aligned}
I_2 &\leqslant \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}(\sum_{i=1}^{q}d_{li}^2\delta_{li})(\sum_{i=1}^{q}e_{ti}^2\delta_{ti})\\
&= \tfrac{1}{n}\textstyle\sum_{l=1}^{n}\sum_{j=2}^{k}(\tfrac{1}{q}\sum_{i\in M_{mj}}d_{li}^2)\mathrm{I}(l\in T_j)(\tfrac{1}{q}\sum_{i\in M_{mu}}e_{ti}^2)\\
&\leqslant O_p(\tfrac{\max_{2\leqslant j\leqslant k}(n_j)}{nn_1})+O(\tfrac{\max_{2\leqslant j\leqslant k}(n_jq_j)}{nq^2}).
\end{aligned}$$

Similarly, we have

$$\begin{aligned} I_3 &\leqslant \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}(\sum_{i=1}^{q} e_{li}^2\delta_{li})(\sum_{i=1}^{q} d_{ti}^2\delta_{ti}) \\ &= \tfrac{1}{n}\textstyle\sum_{l=1}^{n}\sum_{j=2}^{k}(\frac{1}{q}\sum_{i\in M_{mj}} e_{li}^2)\mathrm{I}(l\in T_j)(\frac{1}{q}\sum_{i\in M_{mu}} d_{ti}^2) \\ &\leqslant O_p(\tfrac{\max_{2\leqslant j\leqslant k}(n_j)}{nn_1}) + O(\tfrac{\max_{2\leqslant j\leqslant k}(n_j)q_u}{nq^2}). \end{aligned}$$

Based on Assumption (C3), we have

$$\begin{aligned} I_4 &= \tfrac{1}{nq^2}\textstyle\sum_{l=1}^{n}\sum_{i=1}^{q}\sum_{a=1}^{q} e_{li}e_{la}e_{ti}e_{ta}\delta_{li}\delta_{la}\delta_{ti}\delta_{ta} \\ &= \tfrac{1}{nq}\textstyle\sum_{j=2}^{k}\sum_{l=1}^{n}\{\frac{1}{q}\sum_{i\in M_{mj}\bigcap M_{mu}}\sum_{a\in M_{mj}\bigcap M_{mu}} e_{li}e_{la}e_{ti}e_{ta}\}\mathrm{I}(l\in T_j) \\ &\leqslant O_p(\tfrac{\max_{2\leqslant j\leqslant k}(n_j)}{nq}) \end{aligned}$$

since

$$\begin{aligned} & E\{\tfrac{1}{q}\textstyle\sum_{i\in M_{mj}\bigcap M_{mu}}\sum_{a\in M_{mj}\bigcap M_{mu}} e_{li}e_{la}e_{ti}e_{ta}\} \\ = \; & \tfrac{1}{q}\textstyle\sum_{i\in M_{mj}\bigcap M_{mu}}\sum_{a\in M_{mj}\bigcap M_{mu}} \tau_{ia}^2 \leqslant C \quad \text{if} \quad l\neq t \end{aligned}$$

and

$$\begin{aligned} & E\{\tfrac{1}{q}\textstyle\sum_{i\in M_{mj}\bigcap M_{mu}}\sum_{a\in M_{mj}\bigcap M_{mu}} e_{li}e_{la}e_{ti}e_{ta}\} \\ = \; & \tfrac{1}{q}\textstyle\sum_{i\in M_{mj}\bigcap M_{mu}}\sum_{a\in M_{mj}\bigcap M_{mu}}\{E(e_{ti}^2e_{ta}^2)-\tau_{ia}^2\} \\ & +\tfrac{1}{q}\textstyle\sum_{i\in M_{mj}\bigcap M_{mu}}\sum_{a\in M_{mj}\bigcap M_{mu}} \tau_{ia}^2 \leqslant C \;\text{ if }\; l=t. \end{aligned}$$

Thus, we can obtain $\frac{1}{nq^2}\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^\top\widehat{\mathbf{D}}_t\|^2 = O_p\{\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n\min(n_1,q)}\}$. The proof of lemma is completed.

Lemma 14. *Under Assumptions (C1)-(C5), we have*

$$\frac{1}{n}\sum_{t=1}^{n} d_{ti}^2\delta_{ti} = O_p\left\{\frac{\max_{2\leqslant j\leqslant k}(n_jq_j)}{nq^2}\right\} + O_p\left(\frac{1}{\min(q^2,n_1)}\right) \;\text{ for }\; i=1,\ldots,q.$$

Proof. Based on the definition of d_{ti}, we have

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} d_{ti}^2\delta_{ti} &= \frac{1}{n}\sum_{t=1}^{n}(\widetilde{\boldsymbol{\lambda}}_i^\top \widetilde{\mathbf{w}}_t - \boldsymbol{\lambda}_i^\top \mathbf{w}_t)^2\delta_{ti} \\
&\leqslant C\{\frac{1}{n}\sum_{t=1}^{n}\|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\|\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t\|^2\delta_{ti} \\
&\quad +\frac{1}{n}\sum_{t=1}^{n}\|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\|\mathbf{H}_1\|^2\|\mathbf{w}_t\|^2\delta_{ti} \\
&\quad +\frac{1}{n}\sum_{t=1}^{n}\|\boldsymbol{\lambda}_i\|^2\|\mathbf{H}_1^{-1}\|^2\|\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t\|^2\delta_{ti}\} \\
&= C(I_1 + I_2 + I_3).
\end{aligned}$$

Based on Lemma 8, we have

$$\begin{aligned}
I_1 &= \frac{1}{n}\sum_{j=2}^{k}(\sum_{t\in T_j}\|\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t\|^2)\|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2 \mathrm{I}(i \in M_{mj}) \\
&\leqslant \sum_{j=2}^{k} O_p\{\frac{n_j q_j^2}{nq^2\min(n_1,q_j)\min(n_1,q^2)}\}\mathrm{I}(i \in M_{mj}).
\end{aligned}$$

For Assumption (C1), we have $I_2 \leqslant \|\widetilde{\boldsymbol{\lambda}}_i - \mathbf{H}_1^{-1}\boldsymbol{\lambda}_i\|^2\|\mathbf{H}_1\|^2(\frac{1}{n}\sum_{t=1}^{n}\|\mathbf{w}_t\|^2) = O_p(\frac{1}{\min(n_1,q^2)})$. For Assumption (C2), we have

$$\begin{aligned}
I_3 &= \frac{1}{n}\sum_{j=2}^{k}\sum_{t\in T_j}\|\boldsymbol{\lambda}_i\|^2\|\mathbf{H}_1^{-1}\|^2\|\widetilde{\mathbf{w}}_t - \mathbf{H}_1^\top \mathbf{w}_t\|^2 \mathrm{I}(i \in M_{mj}) \\
&\leqslant \sum_{j=2}^{k} O_p(\frac{n_j q_j^2}{nq^2\min(n_1,q_j)})\mathrm{I}(i \in M_{mj}).
\end{aligned}$$

Thus, we can obtain $\frac{1}{n}\sum_{t=1}^{n} d_{ti}^2\delta_{ti} \leqslant O_p\{\frac{\max_{2\leqslant j\leqslant k}(n_j q_j)}{nq^2}\} + O_p(\frac{1}{\min(q^2,n_1)})$. The proof of lemma is completed.

Proof of Theorem 2. Using the definition of $\widehat{\mathbf{W}}$, we have $\frac{1}{nq}\widehat{\mathbf{Z}}\widehat{\mathbf{Z}}^\top\widehat{\mathbf{W}} = \widehat{\mathbf{W}}\widehat{\mathbf{V}}$ where $\widehat{\mathbf{V}}$ is a diagonal matrix whose diagonal elements are the first r

eigenvalues of $\widehat{\mathbf{Z}}\widehat{\mathbf{Z}}^\top/(nq)$ in decreasing order. We have

$$\begin{aligned}
\widehat{\mathbf{w}}_t &= \tfrac{1}{nq}\widehat{\mathbf{V}}^{-1}\widehat{\mathbf{W}}^\top\widehat{\mathbf{Z}}\widehat{\mathbf{Z}}_t \\
&= \tfrac{1}{nq}\widehat{\mathbf{V}}^{-1}\widehat{\mathbf{W}}^\top(\mathbf{Z}+\widehat{\mathbf{D}})(\mathbf{Z}_t+\widehat{\mathbf{D}}_t) \\
&= \tfrac{1}{nq}\widehat{\mathbf{V}}^{-1}\widehat{\mathbf{W}}^\top(\mathbf{W}\boldsymbol{\Lambda}^\top+\boldsymbol{e}+\widehat{\mathbf{D}})(\boldsymbol{\Lambda}\mathbf{w}_t+\boldsymbol{e}_t+\widehat{\mathbf{D}}_t) \\
&= \tfrac{1}{nq}\widehat{\mathbf{V}}^{-1}\{\widehat{\mathbf{W}}^\top\mathbf{W}\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\mathbf{w}_t+\widehat{\mathbf{W}}^\top\mathbf{W}\boldsymbol{\Lambda}^\top\boldsymbol{e}_t+\widehat{\mathbf{W}}^\top\mathbf{W}\boldsymbol{\Lambda}^\top\widehat{\mathbf{D}}_t+\widehat{\mathbf{W}}^\top\boldsymbol{e}\boldsymbol{\Lambda}\mathbf{w}_t \\
&\quad +\widehat{\mathbf{W}}^\top\boldsymbol{e}\boldsymbol{e}_t+\widehat{\mathbf{W}}^\top\boldsymbol{e}\widehat{\mathbf{D}}_t+\widehat{\mathbf{W}}^\top\widehat{\mathbf{D}}\boldsymbol{\Lambda}\mathbf{w}_t+\widehat{\mathbf{W}}^\top\widehat{\mathbf{D}}\boldsymbol{e}_t+\widehat{\mathbf{W}}^\top\widehat{\mathbf{D}}\widehat{\mathbf{D}}_t\}.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
&\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t \\
=\ &\widehat{\mathbf{V}}^{-1}\{\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\mathbf{W}\boldsymbol{\Lambda}^\top\boldsymbol{e}_t+\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\mathbf{W}\boldsymbol{\Lambda}^\top\widehat{\mathbf{D}}_t+\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\boldsymbol{e}\boldsymbol{\Lambda}\mathbf{w}_t+\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\boldsymbol{e}\boldsymbol{e}_t \\
&+\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\boldsymbol{e}\widehat{\mathbf{D}}_t+\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\widehat{\mathbf{D}}\boldsymbol{\Lambda}\mathbf{w}_t+\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\widehat{\mathbf{D}}\boldsymbol{e}_t+\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\widehat{\mathbf{D}}\widehat{\mathbf{D}}_t\} \\
=\ &\widehat{\mathbf{V}}^{-1}\textstyle\sum_{i=1}^8 I_i.
\end{aligned}$$

Using the definition of $\widehat{\mathbf{D}}_t$, we have

$$\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t=\begin{cases}\widehat{\mathbf{V}}^{-1}(I_1+I_3+I_4+I_6+I_7) & t\in T_1\\ \widehat{\mathbf{V}}^{-1}\sum_{i=1}^8 I_i & t\in T_u,\ u=2,\dots,k.\end{cases}$$

For Assumptions (C1), (C2) and (C3) together with $\frac{\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}}{n}=\mathbf{I}_r$, we have

$$\begin{aligned}
\|I_1\| &= \|\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\mathbf{W}\boldsymbol{\Lambda}^\top\boldsymbol{e}_t\|\leqslant\tfrac{1}{\sqrt{q}}\left\|\tfrac{\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}}{n}\right\|^{1/2}\left\|\tfrac{\mathbf{W}^\top\mathbf{W}}{n}\right\|^{1/2}\|\tfrac{1}{\sqrt{q}}\textstyle\sum_{i=1}^q\boldsymbol{\lambda}_i e_{ti}\| \\
&= O_p(\tfrac{1}{\sqrt{q}})
\end{aligned}$$

where $1/\sqrt{q}\sum_{i=1}^q\boldsymbol{\lambda}_i e_{ti}=O_p(1)$ since $E\{1/\sqrt{q}\sum_{i=1}^q\boldsymbol{\lambda}_i e_{ti}\}=0$ and $E(\|1/\sqrt{q}\sum_{i=1}^q\boldsymbol{\lambda}_i e_{ti}\|^2)\leqslant\bar{\lambda}/q\sum_{i=1}^q\tau_{ii}\leqslant C$. Based on Assumption (C1) and (C3) with $\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}/n = \mathbf{I}_r$, we have

$$\begin{aligned}\|I_3\| &= \|\tfrac{1}{nq}\widehat{\mathbf{W}}^\top \boldsymbol{e}\boldsymbol{\Lambda}\mathbf{w}_t\| = \|\tfrac{1}{nq}\textstyle\sum_{l=1}^n \widehat{\mathbf{w}}_l\mathbf{w}_t^\top\boldsymbol{\Lambda}^\top\boldsymbol{e}_l\| \\ &\leqslant \tfrac{1}{\sqrt{q}}(\tfrac{1}{n}\textstyle\sum_{l=1}^n \|\widehat{\mathbf{w}}_l\|^2)^{1/2}\|\mathbf{w}_t\|\tfrac{1}{n}\textstyle\sum_{l=1}^n\|\tfrac{1}{\sqrt{q}}\textstyle\sum_{i=1}^q \boldsymbol{\lambda}_i e_{li}\|^2)^{1/2} \\ &= O_p(\tfrac{1}{\sqrt{q}}).\end{aligned}$$

Based on Assumption (C3), we can obtain

$$\begin{aligned}\|I_4\| &= \|\tfrac{1}{nq}\widehat{\mathbf{W}}^\top\boldsymbol{e}\boldsymbol{e}_t\| = \|\tfrac{1}{nq}\textstyle\sum_{l=1}^n \widehat{\mathbf{w}}_l\boldsymbol{e}_l^\top\boldsymbol{e}_t\| \\ &\leqslant \tfrac{1}{\sqrt{q}}(\tfrac{1}{n}\textstyle\sum_{l=1}^n\|\widehat{\mathbf{w}}_l\|^2)^{1/2}(\tfrac{1}{nq}\textstyle\sum_{l=1}^n\|\boldsymbol{e}_l^\top\boldsymbol{e}_t\|^2)^{1/2} \leqslant O_p(\tfrac{1}{\sqrt{q}}),\end{aligned}$$

where $\frac{1}{nq}\sum_{l=1}^n\|\boldsymbol{e}_l^\top\boldsymbol{e}_t\|^2 = O_p(1)$ and

$$\begin{aligned}E\{\tfrac{1}{nq}\textstyle\sum_{l=1}^n\|\boldsymbol{e}_l^\top\boldsymbol{e}_t\|^2\} &= \tfrac{1}{nq}\textstyle\sum_{i=1}^q\sum_{j=1}^q\sum_{l=1}^n E(e_{li}e_{lj}e_{ti}e_{tj}) \\ &= \tfrac{1}{nq}\textstyle\sum_{i=1}^q\sum_{j=1}^q\{n\tau_{ij}^2 + E(e_{ti}^2e_{tj}^2) - \tau_{ij}^2\} \\ &= \tfrac{1}{q}\textstyle\sum_{i=1}^q\sum_{j=1}^q\tau_{ij}^2 + \tfrac{1}{nq}\textstyle\sum_{i=1}^q\sum_{j=1}^q\{E(e_{ti}^2e_{tj}^2) - \tau_{ij}^2\} \\ &\leqslant C.\end{aligned}$$

Based on Assumption (C1), $\frac{\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}}{n} = \mathbf{I}_r$ and Lemma 9, we have for $t \in T_u,\ u = 2,\ldots,k$ $\|I_2\| = \|\frac{1}{nq}\widehat{\mathbf{W}}^\top\mathbf{W}\boldsymbol{\Lambda}^\top\widehat{\mathbf{D}}_t\| \leqslant \|\frac{\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}}{n}\|^{1/2}\|\frac{\mathbf{W}^\top\mathbf{W}}{n}\|^{1/2}\|\frac{1}{q}\boldsymbol{\Lambda}^\top\widehat{\mathbf{D}}_t\| = O_p(1)\|\frac{1}{q}\boldsymbol{\Lambda}^\top\widehat{\mathbf{D}}_t\| = O_p(\frac{1}{\min(\sqrt{n_1},\sqrt{q})})$. For $\frac{\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}}{n} = \mathbf{I}_r$ and Lemma 10, we have

$$\begin{aligned}\|I_5\| &\leqslant (\tfrac{1}{n}\textstyle\sum_{l=1}^n\|\widehat{\mathbf{w}}_l\|^2)^{1/2}(\tfrac{1}{nq^2}\textstyle\sum_{l=1}^n\|\boldsymbol{e}_l^\top\widehat{\mathbf{D}}_t\|^2)^{1/2} \\ &= O_p(1)(\tfrac{1}{nq^2}\textstyle\sum_{l=1}^n\|\boldsymbol{e}_l^\top\widehat{\mathbf{D}}_t\|^2)^{1/2} = O_p(\tfrac{1}{\min(\sqrt{n_1},\sqrt{q})})\end{aligned}$$

for $t \in T_u,\ u = 2,\ldots,k$. For Assumption (C1), $\frac{\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}}{n} = \mathbf{I}_r$ and Lemma 11, we have

$$\begin{aligned} \|I_6\| &= \|\tfrac{1}{nq}\textstyle\sum_{l=1}^{n}\widehat{\mathbf{w}}_l\mathbf{w}_t^\top\mathbf{\Lambda}^\top\widehat{\mathbf{D}}_l\| \\ &\leqslant \tfrac{1}{\sqrt{q}}(\tfrac{1}{n}\textstyle\sum_{l=1}^{n}\|\widehat{\mathbf{w}}_l\|^2)^{1/2}\|\mathbf{w}_t^\top\|(\tfrac{1}{nq}\sum_{l=1}^{n}\|\mathbf{\Lambda}^\top\widehat{\mathbf{D}}_l\|^2)^{1/2} \\ &\leqslant O_p\{\tfrac{\max_{2\leqslant j\leqslant k}(\sqrt{n_j})}{\sqrt{nq}}\}. \end{aligned}$$

Similarly, for Lemma 12, we have

$$\|I_7\| \leqslant (\frac{1}{n}\sum_{l=1}^{n}\|\widehat{\mathbf{w}}_l\|^2)^{1/2}(\frac{1}{nq^2}\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^\top\boldsymbol{e}_t\|^2)^{1/2} \leqslant O_p\{\frac{\max_{2\leqslant j\leqslant k}(\sqrt{n_j})}{\sqrt{n}\min(\sqrt{n_1},\sqrt{q})}\}.$$

Based on Lemma 13, we can obtain

$$\|I_8\| \leqslant (\frac{1}{n}\sum_{l=1}^{n}\|\widehat{\mathbf{w}}_l\|^2)^{1/2}(\frac{1}{nq^2}\sum_{l=1}^{n}\|\widehat{\mathbf{D}}_l^\top\widehat{\mathbf{D}}_t\|^2)^{1/2} \leqslant O_p\{\frac{\max_{2\leqslant j\leqslant k}(\sqrt{n_j})}{\sqrt{n}\min(\sqrt{n_1},\sqrt{q})}\}$$

for $t \in T_u,\ u = 2,\ldots,k$. Similarly to the augment of Lemma A1, we can obtain that $\widehat{\mathbf{V}}^{-1}$ is bounded. Thus, we have that

$$\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t = O_p(\frac{1}{\sqrt{q}}) + O_p\{\frac{\max_{2\leqslant j\leqslant k}(\sqrt{n_j})}{\sqrt{nn_1}}\} = O_p(\frac{1}{\min(\sqrt{q},\sqrt{n_1})})$$

for $t \in T_1$, and

$$\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t = O_p(\frac{1}{\min(\sqrt{q},\sqrt{n_1})}) \ \text{ for } \ t \in T_u,\ u = 2,\ldots,k.$$

The proof of the part (i) in theorem 2 is completed.

We prove the part (ii) as the follow. From $\widehat{\mathbf{\Lambda}} = \widehat{\mathbf{Z}}^\top\widehat{\mathbf{W}}/n = \sum_{t=1}^{n}\widehat{\mathbf{Z}}_t\widehat{\mathbf{w}}_t^\top/n$ and $\widehat{\mathbf{Z}}_t = \mathbf{\Lambda}\mathbf{w}_t + \boldsymbol{e}_t + \widehat{\mathbf{D}}_t$, we have $\widehat{\boldsymbol{\lambda}}_i = \frac{1}{n}\sum_{t=1}^{n}\widehat{\mathbf{w}}_t\mathbf{w}_t^\top\boldsymbol{\lambda}_i + \frac{1}{n}\sum_{t=1}^{n}\widehat{\mathbf{w}}_t e_{ti} + \frac{1}{n}\sum_{t=1}^{n}\widehat{\mathbf{w}}_t\widehat{\mathbf{D}}_{ti}$. Writing $\mathbf{w}_t = \mathbf{w}_t - \mathbf{H}^{-1\top}\widehat{\mathbf{w}}_t + \mathbf{H}^{-1\top}\widehat{\mathbf{w}}_t$ and using $\frac{1}{n}\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}} = \mathrm{I}_r$, we have

$$\begin{aligned}\widehat{\boldsymbol{\lambda}}_i - \mathbf{H}^{-1}\boldsymbol{\lambda}_i &= -\tfrac{1}{n}\textstyle\sum_{t=1}^n(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t)(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t)^\top\mathbf{H}^{-1}\boldsymbol{\lambda}_i \\ &\quad -\tfrac{1}{n}\textstyle\sum_{t=1}^n \mathbf{H}^\top\mathbf{w}_t(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t)^\top\mathbf{H}^{-1}\boldsymbol{\lambda}_i \\ &\quad +\tfrac{1}{n}\textstyle\sum_{t=1}^n(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t)e_{ti} + \tfrac{1}{n}\sum_{t=1}^n \mathbf{H}^\top\mathbf{w}_t e_{ti} \\ &\quad +\tfrac{1}{n}\textstyle\sum_{t=1}^n(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t)\widehat{\mathbf{D}}_{ti} + \tfrac{1}{n}\sum_{t=1}^n \mathbf{H}^\top\mathbf{w}_t\widehat{\mathbf{D}}_{ti} \\ &= J_1 + J_2 + J_3 + J_4 + J_5 + J_6.\end{aligned}$$

Assumptions (C1) and (C2) together with $\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}/n = \mathbf{I}_r$ imply that $\|\mathbf{H}\| \leqslant \|\frac{\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}}{q}\|\|\frac{\mathbf{W}^\top\mathbf{W}}{n}\|^{1/2}\|\frac{\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}}{n}\|^{1/2}\|\widehat{\mathbf{V}}^{-1}\| = O_p(1)$. Following the part (i) in Theorem 2, we have

$$\|J_1\| \leqslant \frac{1}{n}\sum_{t=1}^n \|\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t\|^2\|\mathbf{H}^{-1}\|\|\boldsymbol{\lambda}_i\| \leqslant O_p(\frac{1}{\min(q, n_1)}).$$

Based on Assumptions (C1)-(C2) and $\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}}/n = \mathbf{I}_r$, we have

$$\begin{aligned}\|J_2\| &\leqslant \|\mathbf{H}\|(\tfrac{1}{n}\textstyle\sum_{t=1}^n \|\mathbf{w}_t\|^2)^{1/2}(\tfrac{1}{n}\sum_{t=1}^n \|\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t\|^2)^{1/2}\|\mathbf{H}^{-1}\|\|\boldsymbol{\lambda}_i\| \\ &\leqslant O_p(\tfrac{1}{\min(\sqrt{q}, \sqrt{n_1})}).\end{aligned}$$

For Assumption (C3), we have

$$\|J_3\| \leqslant (\frac{1}{n}\sum_{t=1}^n \|\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t\|^2)^{1/2}(\frac{1}{n}\sum_{t=1}^n e_{ti}^2)^{1/2} \leqslant O_p(\frac{1}{\min(\sqrt{q}, \sqrt{n_1})}).$$

For Assumption (C4), we have $\|J_4\| = \|\mathbf{H}\|\|\frac{1}{n}\sum_{t=1}^n \mathbf{w}_t e_{ti}\| = O_p(\frac{1}{\sqrt{n}})$.

Based on the definition of $\widehat{\mathbf{D}}_t$, we have

$$\begin{aligned}\|J_5\| &= \|\tfrac{1}{n}\textstyle\sum_{t=1}^n(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t)(d_{ti} - e_{ti})\delta_{ti}\| \\ &\leqslant (\tfrac{1}{n}\textstyle\sum_{t=1}^n \|\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t\|^2\delta_{ti})^{1/2}(\tfrac{1}{n}\sum_{t=1}^n d_{ti}^2\delta_{ti})^{1/2} \\ &\quad +(\tfrac{1}{n}\textstyle\sum_{t=1}^n \|\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t\|^2\delta_{ti})^{1/2}(\tfrac{1}{n}\sum_{t=1}^n e_{ti}^2\delta_{ti})^{1/2},\end{aligned}$$

where

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\|\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t\|^2\delta_{ti} &= \frac{1}{n}\sum_{j=2}^{k}\sum_{t\in T_j}\|\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t\|^2\mathrm{I}(i\in M_{mj})\\ &\leqslant O_p\{\frac{\max_{2\leqslant j\leqslant k}(n_j)}{n\min(q,n_1)}\}=O_p\{\frac{1}{\min(q,n_1)}\}.\end{aligned}$$

For Assumption (C3), we have $\frac{1}{n}\sum_{t=1}^{n}e_{ti}^2\delta_{ti}=O_p(1)$. Based on Lemma 14, we have $\|J_5\|=O_p\{\frac{1}{\min(\sqrt{q},\sqrt{n_1})}\}$. Similarly, we can prove $\|J_6\|=O_p\{\frac{1}{\min(\sqrt{q},\sqrt{n_1})}\}$. Thus, we can obtain $\widehat{\boldsymbol{\lambda}}_i-\mathbf{H}^{-1}\boldsymbol{\lambda}_i=O_p(\frac{1}{\min(\sqrt{q},\sqrt{n_1})})$. The proof of part (ii) is completed.

Proof of Theorem 3. Based on the definite of $\widehat{\boldsymbol{\alpha}}$, we have

$$\begin{aligned}\widehat{\boldsymbol{\alpha}} &= (\widehat{\mathbf{W}}^\top\widehat{\mathbf{W}})^{-1}\widehat{\mathbf{W}}^\top\mathbf{Y}\\ &= (\sum_{t=1}^{n}\widehat{\mathbf{w}}_t\widehat{\mathbf{w}}_t^\top)^{-1}\sum_{t=1}^{n}\widehat{\mathbf{w}}_t y_t\\ &= \{\sum_{t=1}^{n}(\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t)(\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t)^\top\\ &\quad +\sum_{t=1}^{n}(\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t)\mathbf{w}_t^\top\mathbf{H}+\mathbf{H}^\top\sum_{t=1}^{n}\mathbf{w}_t(\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t)^\top\\ &\quad +\mathbf{H}^\top\sum_{t=1}^{n}\mathbf{w}_t\mathbf{w}_t^\top\mathbf{H}\}^{-1}\{\sum_{t=1}^{n}(\widehat{\mathbf{w}}_t-\mathbf{H}^\top\mathbf{w}_t)y_t+\mathbf{H}^\top\sum_{t=1}^{n}\mathbf{w}_t y_t\}\\ &= \{\mathbf{H}^\top\sum_{t=1}^{n}\mathbf{w}_t\mathbf{w}_t^\top\mathbf{H}+o_p(1)\}^{-1}\\ &\quad \times\{\mathbf{H}^\top\sum_{t=1}^{n}\mathbf{w}_t\mathbf{w}_t^\top\boldsymbol{\alpha}+\mathbf{H}^\top\sum_{t=1}^{n}\mathbf{w}_t\varepsilon_t+o_p(1)\}\\ &= \mathbf{H}^{-1}\boldsymbol{\alpha}+\mathbf{H}^{-1}(\sum_{t=1}^{n}\mathbf{w}_t\mathbf{w}_t^\top)^{-1}\sum_{t=1}^{n}\mathbf{w}_t\varepsilon_t+o_p(1)\end{aligned}$$

since $\|\mathbf{H}\|=O_p(1)$ and Assumption (C1) together with Theorem 2 (i) holds. Thus, from Assumption (C6), we can obtain

$$\widehat{\boldsymbol{\alpha}}-\mathbf{H}^{-1}\boldsymbol{\alpha}=\mathbf{H}^{-1}(\sum_{t=1}^{n}\mathbf{w}_t\mathbf{w}_t^\top)^{-1}\sum_{t=1}^{n}\mathbf{w}_t\varepsilon_t+o_p(1)\xrightarrow{\mathrm{p}}0.$$

The proof of part (i) is completed. Similarly, we have

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}}^\top\widehat{\mathbf{w}}_t - \boldsymbol{\alpha}^\top\mathbf{w}_t &= \widehat{\boldsymbol{\alpha}}^\top\widehat{\mathbf{w}}_t - (\mathbf{H}^{-1}\boldsymbol{\alpha})^\top(\mathbf{H}^\top\mathbf{w}_t) \\
&= (\widehat{\boldsymbol{\alpha}} - \mathbf{H}^{-1}\boldsymbol{\alpha})^\top(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t) + (\widehat{\boldsymbol{\alpha}} - \mathbf{H}^{-1}\boldsymbol{\alpha})^\top\mathbf{H}^\top\mathbf{w}_t \\
&\quad + (\mathbf{H}^{-1}\boldsymbol{\alpha})^\top(\widehat{\mathbf{w}}_t - \mathbf{H}^\top\mathbf{w}_t) \\
&\overset{\mathrm{p}}{\to} 0.
\end{aligned}$$

The proof of part (ii) is completed.

Bibliography

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

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

Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society B 70*, pp. 849–911.