Supplementary Appendix to "An Improved Inference for IV Regressions"*

Liyu Dou[a], Pengjin Min[a], Wenjie Wang[b], and Yichong Zhang[a]

[a]*Singapore Management University, Singapore*

[b]*Nanyang Technological University, Singapore*

February 6, 2026

A Combination Test with Weak Low-Dimensional IVs

In this section, we study the property of our combination test when Assumption 2 is violated. Specifically, we assume that the identification strength provided by the low-dimensional IVs is weak, rendering the Wald test invalid. On the other hand, we assume that the identification strength provided by the many IVs is strong. To characterize the limiting behavior of the low-dimensional IVs under weak identification, we introduce the following assumptions.

Assumption 2′. *Let Assumption 2.1 hold. In addition, the following conditions hold almost surely:*

1. *For* $r_n = \left\|z^\top \Pi\right\|_2$, $r_n/\sqrt{n} = O(1)$;

*We are grateful to Michal Kolésar for helpful comments. Any and all errors are our own.

2. It holds that

$$\frac{1}{n}\begin{pmatrix} \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{e}_{[g]}\right)\left(z_{[g]}^\top \tilde{e}_{[g]}\right)^\top & \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{e}_{[g]}\right)\left(z_{[g]}^\top \tilde{V}_{[g]}\right)^\top \\ \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{V}_{[g]}\right)\left(z_{[g]}^\top \tilde{e}_{[g]}\right)^\top & \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{V}_{[g]}\right)\left(z_{[g]}^\top \tilde{V}_{[g]}\right)^\top \end{pmatrix} > 0,$$

in the matrix sense for all n large enough.

Remark A.1. Compared to Assumption 2, the main difference here is that we have $r_n/\sqrt{n} = O(1)$. This corresponds to weak identification of the parameter of interest β under the low-dimensional IVs, since in this case the deterministic part and the random part of $z^\top X$ are of the same order. This setting is similar to the weak-IV asymptotics considered in Staiger and Stock (1997), where $z^\top X/\sqrt{n}$ converges to a random limit instead of diverging to infinity (the latter would happen under a standard asymptotics where $z^\top X/n$ is assumed to converge to a non-zero fixed limit).

In the following, we first present the results for the case with $d_z = 1$. Note that it is the most important case for empirical applications of IV regressions. For instance, 101 out of 230 specifications in Andrews, Stock, and Sun (2019)'s sample and 1,087 out of 1,359 in Young (2022)'s sample feature one endogenous regressor and one IV. Similarly, Lee, McCrary, Moreira, and Porter (2022) find that 61 out of 123 IV papers published in *AER* between 2013 and 2019 use one endogenous regressor and one IV. For these applications, empirical researchers can generate many IVs by using polynomials or interactions based on their one-dimensional base IV and control variables. Then, it is possible to achieve efficiency improvement using our combination procedure. Furthermore, d_z is also equal to one in the widely used shift-share IV regressions. It turns out that in this case, the local asymptotic power function of our combination test is equal to that of the asymptotically optimal test based on $z^\top e(\beta_0)/\sqrt{\Omega}$, $LM(\beta_0)$, and AR, where the first statistic corresponds to the conventional cluster-robust AR test using z as instruments.

Theorem A.1. *Suppose that $d_z = 1$. Assume that the following limit exist (almost surely):*

$$\bar{\rho}_1 = \lim_{n\to\infty} \frac{1}{\sqrt{\Omega\Sigma}} \sum_{g\in[G]} \mathbb{E}\left[\left(z_{[g]}^\top \tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)\right],$$
$$\rho_2 = \lim_{n\to\infty} \frac{2}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \mathbb{E}\left[\left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\right],$$

and $\bar{\rho}_1^2 + \rho_2^2 < 1$. Under Assumptions 1, 2 and 3, and assume that $\Pi^\top\Pi/\sqrt{K} \to \infty$, then we have:

1. *Suppose that there exists a deterministic sequence $d_n \downarrow 0$ such that*

$$d_n \Phi_2^{-1/2} \to a > 0, \quad \text{and} \quad \beta - \beta_0 = \delta d_n,$$

for some fixed δ, then

$$\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] = \mathbb{P}\left(\left(-\frac{\bar{\rho}_1}{\sqrt{1-\bar{\rho}_1^2-\rho_2^2}}\mathcal{N}_1 + \frac{1}{\sqrt{1-\bar{\rho}_1^2-\rho_2^2}}\mathcal{N}_2 - \frac{\rho_2}{\sqrt{1-\bar{\rho}_1^2-\rho_2^2}}\mathcal{N}_3\right)^2 \geqslant \mathbb{C}_\alpha\right)$$
$$= \mathbb{P}\left(\chi_1^2\left(\delta^2 \frac{a^2}{1-\bar{\rho}_1^2-\rho_2^2}\right) \geqslant \mathbb{C}_\alpha\right),$$

where
$$\begin{pmatrix}\mathcal{N}_1\\ \mathcal{N}_2\\ \mathcal{N}_3\end{pmatrix} \overset{d}{=} \mathcal{N}\left(\begin{pmatrix}0\\ a\delta\\ 0\end{pmatrix}, \begin{pmatrix}1 & \bar{\rho}_1 & 0\\ \bar{\rho}_1 & 1 & \rho_2\\ 0 & \rho_2 & 1\end{pmatrix}\right).$$

2. *Suppose that $\beta - \beta_0 = \delta$ for some fixed $\delta \neq 0$, then $\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] = 1$.*

Theorem A.1 implies that, in the scalar case ($d_z = 1$), our approach is equivalent to an optimal combination of the low-dimensional AR, Jackknife LM, and Jackknife AR. A key scalar-specific feature is that the discrepancy between the low-dimensional AR and low-dimensional Wald (i.e., $T(\beta_0)$) effectively reduces to a random sign that cannot be consistently estimated when the IV z is weak; correspondingly, $\hat{\rho}_1$ is inconsistent, and the low-dimensional Wald is non-normal. However, by appropriately choosing combination weights, this unidentifiable sign component cancels out, preserving the validity of our combination test in the important scalar setting.

When $d_z > 1$, it is no longer possible to cancel the random sign. Nevertheless, if the correlation between $T(\beta_0)$ and $LM(\beta_0)$ is asymptotically negligible, our combined test remains unaffected by $T(\beta_0)$, as shown in the next result.

Theorem A.2. *Assume that the following limit exists (almost surely)*

$$\rho = \lim_{n\to\infty} \frac{2}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \mathbb{E}\left[\left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\right],$$

with $\rho^2 < 1$. *Under Assumptions 1, 2' and 3, and assume that* $\Pi^\top\Pi/\sqrt{K} \to \infty$. *If further assume* $\Pi^\top\Pi/K \to 0$, *then we have:*

1. *Suppose there exists a deterministic sequence* $d_n \downarrow 0$ *such that*

$$d_n \Phi_2^{-1/2} \to a > 0, \quad \text{and} \quad \beta - \beta_0 = \delta d_n,$$

for some fixed δ, *then*

$$\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] = \mathbb{P}\left(\left(\frac{1}{\sqrt{1-\rho^2}}\mathcal{N}_1 - \frac{\rho}{\sqrt{1-\rho^2}}\mathcal{N}_2\right)^2 \geqslant \mathbb{C}_\alpha\right) = \mathbb{P}\left(\chi_1^2\left(\delta^2\frac{a^2}{1-\rho^2}\right) \geqslant \mathbb{C}_\alpha\right),$$

where $\begin{pmatrix}\mathcal{N}_1\\ \mathcal{N}_2\end{pmatrix} \overset{d}{=} \mathcal{N}\left(\begin{pmatrix}a\delta\\ 0\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right).$

2. *Suppose that* $\beta - \beta_0 = \delta$ *for some fixed* $\delta \neq 0$, *then* $\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] = 1$.

Under the local alternative $\beta - \beta_0 = \delta d_n$, $\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right]$ coincides exactly with the local asymptotic power function of the asymptotically optimal test based on $LM(\beta_0)$ and AR in Lim, Wang, and Zhang (2024). The assumption that $\Pi^\top\Pi/K \to 0$ is similar to the assumption in Mikusheva and Sun (2022, Theorem 4). As pointed out by Mikusheva and Sun (2022), this condition is quite weak as it still covers both weakly and strongly identified cases (with many IVs). In addition, we notice that under fixed alternatives, the combination test remains consistent even when the identification strength of the low-dimensional IVs is weak.

Note also that it is possible to combine the low-dimensional AR, Jackknife LM, and Jackknife AR directly, provided that the many IVs are strong. In this way, the combination test is robust against weak low-dimensional IVs when $d_z > 1$ without further assumptions. In this paper, however, we emphasize Wald-based inference for the low-dimensional IV setting because it is the workhorse in empirical applications and, when low-dimensional IVs is strong, Wald is more powerful than AR in the overidentified case.

Finally, if both low-dimensional IVs and many IVs are weak, then by using null-imposed variance estimators, the combination test that combines the low-dimensional AR, Jackknife LM, and Jackknife AR is robust to arbitrary weak identification, regardless of d_z provided that it is bounded. It is also possible to consider a broader combination that includes weak-identification-robust AR and LM (along with Jackknife LM and Jackknife AR). This is beyond the scope of the current paper.

B Technical Lemmas

We use the following notation throughout Sections B-D. Recall the definition of $\tilde{Y}$, $\tilde{X}$, $\tilde{Z}$, $\tilde{z}$, W, $\tilde{\Pi}$, $\tilde{V}$ and $\tilde{e}$. Denote P_W as the projection matrix of W, $M_W = I_n - P_W$, and let $Y = M_W\tilde{Y}$, $X = M_W\tilde{X}$, $Z = M_W\tilde{Z}$, $z = M_W\tilde{z}$, $\Pi = M_W\tilde{\Pi}$, $V = M_W\tilde{V}$ and $e = M_W\tilde{e}$. Denote P as the projection matrix of Z, $M = I_n - P$, and $Q = M_W(P - \bar{P})M_W$, where $\bar{P}$ is the block diagonal matrix corresponding to P such that the g-th block on its diagonal is $P_{[g,g]}$; also denote $\bar{Q}$ as the block diagonal matrix corresponding to Q. Let $\acute{\Pi} = zA_n z^\top X$, $\hat{\Pi} = M_W(P - \bar{P})\Pi = Q\tilde{\Pi}$, $\bar{\Pi} = (Q - \bar{Q})\tilde{\Pi}$, $\acute{X} = z\hat{A}_n z^\top X$ and $\hat{X} = M_W(P - \bar{P})X$. Finally, we use $\Omega_{\tilde{e}}$ to denote the block diagonal matrix with g-th block $\Omega_g^{\tilde{e},\tilde{e}}$, and $\Omega_{\tilde{V}}$, $\Omega_{\tilde{e},\tilde{V}}$ and $\Omega_{\tilde{V},\tilde{e}}$ are defined similarly.

Lemma B.1. *Under Assumption 1, we have*

$$\max_{g\in[G]}\left(\tilde{\Pi}_{[g]}^\top\tilde{\Pi}_{[g]}\right)^2+\max_{g\in[G]}\mathbb{E}\left(\tilde{V}_{[g]}^\top\tilde{V}_{[g]}\right)^2+\max_{g\in[G]}\mathbb{E}\left(\tilde{e}_{[g]}^\top\tilde{e}_{[g]}\right)^2\leqslant C,$$
$$\max_{g\in[G]}\left(\Pi_{[g]}^\top\Pi_{[g]}\right)^2+\max_{g\in[G]}\mathbb{E}\left(V_{[g]}^\top V_{[g]}\right)^2+\max_{g\in[G]}\mathbb{E}\left(e_{[g]}^\top e_{[g]}\right)^2\leqslant C,$$

for some constant $C<\infty$. In addition, let $\hat{\gamma}_{\tilde{e}}=(W^\top W)^{-1}W^\top\tilde{e}$ and $\hat{\gamma}_{\tilde{V}}=(W^\top W)^{-1}W^\top\tilde{V}$, we have

$$\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2=o_P(1),\quad\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{V}}\right\|_2=o_P(1).$$

Lemma B.2. *If Assumptions 1 and 3 hold, then both P and Q are symmetric, and satisfy*

$$\|P\|_{op}=O(1),\quad\left\|\bar{P}\right\|_{op}=O(1),\quad\|P\|_F=O(\sqrt{K}),$$
$$\|Q\|_{op}=O(1),\quad\left\|\bar{Q}\right\|_{op}=o(1),\quad\|Q\|_F=O(\sqrt{K}).$$

In addition, let $\tilde{P}$ be the block lower triangular matrix corresponding to $P-\bar{P}$ (i.e. $\tilde{P}_{[g,h]}=P_{[g,h]}$ for $g>h$ and $\tilde{P}_{[g,h]}=0_{n_g\times n_h}$ otherwise), then

$$\left\|\tilde{P}\tilde{P}^\top\right\|_F=O(\sqrt{K}).$$

Lemma B.3. *Under Assumptions 1 and 3, we have*

$$\tilde{u}^\top\bar{P}P_W\tilde{v}=O_P(1),\quad\tilde{u}^\top P_W\bar{P}\tilde{v}=O_P(1),\quad\tilde{u}^\top\bar{P}P_W\bar{P}\tilde{v}=O_P(1),$$

for $(\tilde{u},\tilde{v})\in\{\tilde{V},\tilde{e}\}\times\{\tilde{V},\tilde{e}\}$.

Lemma B.4. *Under Assumptions 1 and 3, we have*

$$\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g} u_{[h]}^\top P_{[h,g]} v_{[g]}\right)^2 = \frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g} \tilde{u}_{[h]}^\top P_{[h,g]} \tilde{v}_{[g]}\right)^2 + o_P(1)$$

$$= \frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g} \tilde{u}_{[h]}^\top P_{[h,g]} \tilde{v}_{[g]}\right)^2 + o_P(1),$$

$$\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(u_{[h]}^\top P_{[h,g]} v_{[g]}\right)\left(u_{[g]}^\top P_{[g,h]} v_{[h]}\right) = \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{u}_{[h]}^\top P_{[h,g]} \tilde{v}_{[g]}\right)\left(\tilde{u}_{[g]}^\top P_{[g,h]} \tilde{v}_{[h]}\right) + o_P(1)$$

$$= \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{u}_{[h]}^\top P_{[h,g]} \tilde{v}_{[g]}\right)\left(\tilde{u}_{[g]}^\top P_{[g,h]} \tilde{v}_{[h]}\right) + o_P(1),$$

for $(u, v) \in \{V, e\} \times \{V, e\}$, *and the same results hold if we replace* P *with* Q. *In addition, we have*

$$\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{u}_{[h]}^\top P_{[h,g]} \tilde{v}_{[g]}\right)^2 = \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{u}_{[h]}^\top Q_{[h,g]} \tilde{v}_{[g]}\right)^2 + o(1),$$

$$\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{u}_{[h]}^\top P_{[h,g]} \tilde{v}_{[g]}\right)\left(\tilde{u}_{[g]}^\top P_{[g,h]} \tilde{v}_{[h]}\right)$$

$$= \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{u}_{[h]}^\top Q_{[h,g]} \tilde{v}_{[g]}\right)\left(\tilde{u}_{[g]}^\top Q_{[g,h]} \tilde{v}_{[h]}\right) + o(1),$$

for $(\tilde{u}, \tilde{v}) \in \{\tilde{V}, \tilde{e}\} \times \{\tilde{V}, \tilde{e}\}$.

Lemma B.5. *Under Assumptions 1 and 3, we have* $\Sigma \geqslant C(\Pi^\top\Pi + K)$ *for some constant* $C > 0$, *and*

$$\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]}\right)^2$$

$$= \frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]}\right)^2 + o_P(1),$$

$$= \frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\bar{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)^2 + \frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]}\right)^2 + o_P(1),$$

$$
\begin{aligned}
&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right)^2\\
&=\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right)^2\\
&=\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)^2+\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\bar{\Pi}_{[g]}^{\top}\tilde{V}_{[g]}\right)^2\\
&+\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)^2\\
&+\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{V}_{[g]}\right)^2+o_P(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}e_{[g]}\right)\\
&=\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\bar{\Pi}_{[g]}^{\top}\tilde{V}_{[g]}\right)\left(\bar{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\\
&+\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{V}_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{V}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right)+o_P(1).
\end{aligned}
$$

Lemma B.6. *Under Assumptions 1 and 3, we have*

$$
\begin{aligned}
&\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)\\
&=\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^{\top}Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)+o_P(1),
\end{aligned}
$$

$$\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}e_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}e_{[g]}\right)$$
$$=\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)+o_P(1),$$

$$\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}X_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}X_{[g]}\right)$$
$$=\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)$$
$$+\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)$$
$$+\frac{2}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)+o_P(1),$$

and

$$\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}X_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}e_{[g]}\right)$$
$$=\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)$$
$$+\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)+o_P(1).$$

Lemma B.7. *Let $\acute{\beta}$ be a generic estimator of β. Further define*

$$\acute{\Psi}=X^\top z\hat{A}_n\acute{\Omega}\hat{A}_nz^\top X,$$
$$\acute{\Sigma}=\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\acute{e}_{[g]}\right)^2+\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\acute{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\acute{e}_{[g]}\right),$$
$$\acute{\Upsilon}=2\sum_{g,h\in[G]^2,g\neq h}\left(\acute{e}_{[g]}^\top P_{[g,h]}\acute{e}_{[h]}\right)^2,$$

where $\acute{e} = Y - X\acute{\beta}$ *and*

$$\acute{\Omega} = \sum_{g \in [G]} \left(z_{[g]}^\top \acute{e}_{[g]}\right) \left(z_{[g]}^\top \acute{e}_{[g]}\right)^\top .$$

Suppose that $\acute{\beta} \xrightarrow{p} \beta$, *then the following holds.*

1. *If Assumption 1 holds, then*

$$\hat{\Omega}^{-1/2}\Omega^{1/2} = I_{d_z} + o_P(1)$$

2. *If Assumptions 1 and 2 hold, then*

$$\frac{\acute{\Psi}}{\Psi} = 1 + o_P(1).$$

3. *If Assumptions 1 and 3 hold, then*

$$\frac{\acute{\Sigma}}{\Sigma} = 1 + o_P(1),$$
$$\frac{\acute{\Upsilon}}{\Upsilon} = 1 + o_P(1).$$

Lemma B.8. *Under Assumptions 1-3, we have* $\hat{\beta} \xrightarrow{p} \beta$ *and* $(\hat{\beta} - \beta)^2 \Pi^\top \Pi / \sqrt{K} = o_P(1)$. *Alternatively, if Assumptions 1, 2' and 3 hold and* $\Pi^\top \Pi / \sqrt{K} \to \infty$, *we have* $\hat{\beta} \xrightarrow{p} \beta$ *and* $(\hat{\beta} - \beta)^2 \Pi^\top \Pi / \sqrt{K} = o_P(1)$.

Lemma B.9. *Under Assumptions 1-4, if the assumptions for* a_1 *and* a_2 *in Theorem 4.1 hold, then*

$$\begin{pmatrix} T(\beta_0) \\ LM(\beta_0) \\ AR \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{\Psi}} \sum_{g=1}^{G} \acute{\Pi}_{[g]}^\top \tilde{e}_{[g]} \\ \frac{1}{\sqrt{\Sigma}} \left(\sum_{g \in [G]} \hat{\Pi}_{[g]}^\top \tilde{e}_{[g]} + \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]} \right) \\ \frac{1}{\sqrt{\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]} \end{pmatrix} + \begin{pmatrix} a_1 \delta \\ a_2 \delta \\ 0 \end{pmatrix} + o_P(1).$$

Lemma B.10. *If Assumptions 1-4 hold, then*

$$\begin{pmatrix} \frac{1}{\sqrt{\Psi}} \sum_{g=1}^{G} \acute{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} \\ \frac{1}{\sqrt{\Sigma}} \left(\sum_{g \in [G]} \hat{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} + \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \right) \\ \frac{1}{\sqrt{\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \tilde{e}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \end{pmatrix} \rightsquigarrow \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & 0 \\ \rho_1 & 1 & \rho_2 \\ 0 & \rho_2 & 1 \end{pmatrix} \right).$$

Lemma B.11. *Under Assumptions 1 and 3, we have*

$$\frac{1}{\sqrt{\Sigma}} \Omega^{-1/2} \sum_{g \in [G]} \left[\left(z_{[g]}^{\top} \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^{\top} \hat{e}_{[g]} \right) \right] = \frac{1}{\sqrt{\Sigma}} \Omega^{-1/2} \sum_{g \in [G]} \mathbb{E} \left[\left(z_{[g]}^{\top} \tilde{e}_{[g]} \right) \left(\hat{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} \right) \right] + o_P(1),$$

and

$$\begin{aligned} &\frac{1}{\sqrt{\Sigma \Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \left(X_{[g]}^{\top} P_{[g,h]} \hat{e}_{[h]} \right) \left(\hat{e}_{[g]}^{\top} P_{[g,h]} \hat{e}_{[h]} \right) \\ &= \frac{1}{\sqrt{\Sigma \Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{e}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \right) + o_P(1). \end{aligned}$$

Lemma B.12. *Under Assumptions 1–4, we have*

$$\hat{\rho}_1 \xrightarrow{p} \rho_1, \quad \hat{\rho}_2 \xrightarrow{p} \rho_2.$$

If in addition the assumptions for a_1 and a_2 in Theorem 4.1 hold, then

$$\hat{\alpha}_1 \xrightarrow{p} \alpha_1, \quad \hat{\alpha}_2 \xrightarrow{p} \alpha_2.$$

C Proofs of Technical Lemmas

C.1 Proof of Lemma B.1

The first result follows readily from Assumption 1. For the second result, note that for any $g \in [G]$, we have

$$\left(\Pi_{[g]}^{\top}\Pi_{[g]}\right)^2 = \left(\sum_{i\in I_g}\Pi_{i,g}^2\right)^2 \leqslant C\sum_{i\in I_g}\Pi_{i,g}^4 \leqslant C.$$

In addition, by an abuse of notation, for any $i \in [n]$ and $g \in [G]$, we denote $M_W^{(i)}$ as the i-th column of M_W, and $M_{W,[g]}^{(i)} \in \Re^{n_g}$ is vector that collects all elements in the n-dimensional vector $M_W^{(i)}$ that belong to the g-th cluster. Then, we have

$$\begin{aligned}
\mathbb{E}V_{i,g}^4 &= \mathbb{E}\left(\sum_{g\in[G]} M_{W,[g]}^{(i),\top}\tilde{V}_{[g]}\right)^4 \\
&\leqslant C\left(\sum_{g\in[G]}\left(M_{W,[g]}^{(i),\top}M_{W,[g]}^{(i)}\right)^2 + \sum_{g,h\in[G]^2,h\neq g}\left(M_{W,[g]}^{(i),\top}M_{W,[g]}\right)\left(M_{W,[h]}^{(i),\top}M_{W,[h]}\right)\right) \\
&\leqslant C,
\end{aligned}$$

by the first result and the fact that

$$\sum_{g\in[G]}\left(M_{W,[g]}^{(i),\top}M_{W,[g]}^{(i)}\right) = M_W^{(i),\top}M_W^{(i)} = M_{W,ii} \leqslant C.$$

It follows that

$$\mathbb{E}\left(V_{[g]}^{\top}V_{[g]}\right)^2 = \mathbb{E}\left(\sum_{i\in I_g}V_{i,g}^2\right)^2 \leqslant C\sum_{i\in I_g}\mathbb{E}V_{i,g}^4 \leqslant C.$$

Using the same argument, we also have

$$\mathbb{E}\left(e_{[g]}^{\top}e_{[g]}\right)^{2} \leqslant C,$$

and the desired result follows. Finally, for the last result, by Assumption 1, we have $\hat{\gamma}_{\tilde{e}} = O_P(1/\sqrt{n})$ and $\hat{\gamma}_{\tilde{V}} = O_P(1/\sqrt{n})$, and thus

$$\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2^2 \leqslant \max_{1\leqslant g\leqslant G} n_g \times \max_{i\in I_g, g\in[G]}\|W_{i,g}\|_2^2 \times \|\hat{\gamma}_{\tilde{e}}\|_2^2 = o_P(1),$$
$$\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{V}}\right\|_2^2 \leqslant \max_{1\leqslant g\leqslant G} n_g \times \max_{i\in I_g, g\in[G]}\|W_{i,g}\|_2^2 \times \|\hat{\gamma}_{\tilde{V}}\|_2^2 = o_P(1).$$

□

C.2 Proof of Lemma B.2

For the first part of Lemma B.2, the results for P are standard for projection matrix, so we focus on the results for Q. We have

$$\|Q\|_{op} = \left\|M_W(P-\bar{P})M_W\right\|_{op} \leqslant \|M_W\|_{op}^2\left\|P-\bar{P}\right\|_{op} = O(1),$$

and

$$\|Q\|_F = \sqrt{\text{trace}(M_W(P-\bar{P})M_W(P-\bar{P})M_W)} \leqslant C\left\|P-\bar{P}\right\|_F = O(\sqrt{K}).$$

In addition, we note that

$$\bar{Q} = \bar{P}_W\bar{P} + \bar{P}\bar{P}_W - \hat{P},$$

where $\bar{P}_W$ is the block diagonal matrix corresponding to P_W and $\hat{P}$ is a block diagonal matrix such that the g-th block on its diagonal is $\sum_{h=1}^{G} P_{W,[g,h]}P_{[h,h]}P_{W,[h,g]}$ (corresponding to the block diagonals of $P_W \bar{P} P_W$). By Assumption 1, we have

$$\max_{1 \leqslant i \leqslant n} P_{W,ii} \leqslant \max_{i \in I_g, g \in [G]} \|W_{i,g}\|_2^2 \times \lambda_{\max}\left((W^\top W)^{-1}\right) = o(1),$$

and thus

$$\max_{1 \leqslant g \leqslant G} \lambda_{\max}\left(P_{W,[g,g]}\right) \leqslant \max_{1 \leqslant g \leqslant G} n_g \times \max_{1 \leqslant i \leqslant n} P_{W,ii} = o(1),$$

which implies that $\lambda_{\max}\left(\bar{P}_W\right) = o(1)$. It follows that

$$\begin{aligned}
\left\|\bar{Q}\right\|_{op} &\leqslant \left\|\bar{P}_W \bar{P}\right\|_{op} + \left\|\bar{P}\bar{P}_W\right\|_{op} + \left\|\hat{P}\right\|_{op} \\
&\leqslant 2\left\|\bar{P}_W\right\|_{op}\left\|\bar{P}\right\|_{op} + \max_{1 \leqslant g \leqslant G}\left\|\sum_{h \in [G]} P_{W,[g,h]} P_{[h,h]} P_{W,[h,g]}\right\|_{op} \\
&\leqslant C \max_{1 \leqslant g,h \leqslant G}\left\|P_{W,[g,h]} P_{W,[h,g]}\right\|_{op} + o(1) \\
&\leqslant C \max_{1 \leqslant g \leqslant G}\left\|P_{W,[g,g]}\right\|_{op} + o(1) \\
&= o(1).
\end{aligned}$$

For the second part of Lemma B.2, we shall use an argument similar to Chao, Swanson, Hausman, Newey, and Woutersen (2012). A closer inspection of their proof suggests that, all the equalities in the proof of their Lemma B.2. remain unchanged if we replace P_{ij} with $P_{[g,h]}$ and keep the trace operator; note also that we have $\operatorname{trace}\left((P - \bar{P})^4\right) = O(K)$ so that (i) of Lemma B.2. still holds. To obtain (iii) of Lemma B.2. we establish results similar to their Lemma B.1.: for any subset $\mathcal{I}_2$ of the set $\{g,h\}_{g,h=1}^G$, we have

$$\begin{aligned}
\operatorname{trace}\left(\sum_{\mathcal{I}_2} P_{[g,h]} P_{[h,g]} P_{[g,h]} P_{[h,g]}\right) &\leqslant C \sum_{\mathcal{I}_2} \operatorname{trace}\left(P_{[g,h]} P_{[h,g]}\right) \\
&\leqslant C \sum_{g,h \in [G]^2} \operatorname{trace}\left(P_{[g,h]} P_{[h,g]}\right)
\end{aligned}$$

$$= O(K),$$

and similarly for any subset $\mathcal{I}_3$ of the set $\{g,h,k\}_{g,h,k=1}^{G}$, we have

$$\text{trace}\left(\sum_{\mathcal{I}_3} P_{[g,h]}P_{[h,k]}P_{[k,h]}P_{[h,g]}\right) = O(K),$$
$$\text{trace}\left(\sum_{\mathcal{I}_3} P_{[g,h]}P_{[h,g]}P_{[g,k]}P_{[k,g]}\right) = O(K),$$

and then it is easy to see that (iii) of Lemma B.2. holds. To obtain (ii) of Lemma B.2. we define, as in their paper, the following random variables

$$\Delta_1 = \sum_{g<h<k} \left(\xi_{[h]}^\top P_{[h,g]}P_{[g,k]}\xi_{[k]} + \xi_{[g]}^\top P_{[g,h]}P_{[h,k]}\xi_{[k]} + \xi_{[g]}^\top P_{[g,k]}P_{[k,h]}\xi_{[h]}\right)$$
$$\Delta_2 = \sum_{g<h<k} \left(\xi_{[h]}^\top P_{[h,g]}P_{[g,k]}\xi_{[k]} + \xi_{[g]}^\top P_{[g,h]}P_{[h,k]}\xi_{[k]}\right)$$
$$\Delta_3 = \sum_{g<h<k} \left(\xi_{[g]}^\top P_{[g,k]}P_{[k,h]}\xi_{[h]}\right)$$

where $\{\xi_i\}_{i=1}^n$ is a sequence of i.i.d. random variables with mean 0 and variance 1, and independent of $\tilde{Z}, W$ (note that ξ_i are not only independent across clusters but also within clusters), and then it is straightforward to verify that (ii) of Lemma B.2. also holds. These results, together with a similar argument as in the proof of their Lemma B.3., allow us to conclude that $\left\|\tilde{P}\tilde{P}^\top\right\|_F = O(\sqrt{K})$. This concludes the proof. □

C.3 Proof of Lemma B.3

We focus on the case when $\tilde{u} = \tilde{V}$ and $\tilde{v} = \tilde{e}$. We have

$$\left|\tilde{V}^\top P_W \bar{P}\tilde{e}\right| \leqslant \sqrt{\tilde{V}^\top P_W \tilde{V}} \times \sqrt{\tilde{e}^\top \bar{P} P_W \bar{P}\tilde{e}} = O_P(1),$$

because

$$
\begin{aligned}
\mathbb{E}\left(\tilde{V}^\top P_W \tilde{V}\right) &= \text{trace}\left(P_W \Omega_{\tilde{V}}\right) \leqslant \lambda_{\max}(\Omega_{\tilde{V}}) \text{trace}\left(P_W\right) \leqslant C, \\
\mathbb{E}\left(\tilde{e}^\top \bar{P} P_W \bar{P} \tilde{e}\right) &= \text{trace}\left(\bar{P} P_W \bar{P} \Omega_{\tilde{e}}\right) \leqslant \lambda_{\max}(\Omega_{\tilde{e}}) \lambda_{\max}(\bar{P})^2 \text{trace}\left(P_W\right) \leqslant C,
\end{aligned}
$$

since d_w is fixed. The other two terms can be handled similarly. □

C.4 Proof of Lemma B.4

For the first result, we focus on the case when $u = V$ and $v = e$. To show

$$
\frac{1}{K} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} V_{[h]}^\top P_{[h,g]} e_{[g]} \right)^2 - \frac{1}{K} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top P_{[h,g]} \tilde{e}_{[g]} \right)^2 = o_P(1), \tag{C.1}
$$

we note that

$$
\begin{aligned}
& \frac{1}{K} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} V_{[h]}^\top P_{[h,g]} e_{[g]} \right)^2 \\
&= \frac{1}{K} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} (\tilde{V}_{[h]} - W_{[h]} \hat{\gamma}_{\tilde{V}})^\top P_{[h,g]} (\tilde{e}_{[g]} - W_{[g]} \hat{\gamma}_{\tilde{e}}) \right)^2 \\
&= \frac{1}{K} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \left(\begin{array}{c} \tilde{V}_{[h]} P_{[h,g]} \tilde{e}_{[g]} - (W_{[h]} \hat{\gamma}_{\tilde{V}})^\top P_{[h,g]} \tilde{e}_{[g]} \\ - \tilde{V}_{[h]}^\top P_{[h,g]} (W_{[g]} \hat{\gamma}_{\tilde{e}}) + (W_{[h]} \hat{\gamma}_{\tilde{V}})^\top P_{[h,g]} (W_{[g]} \hat{\gamma}_{\tilde{e}}) \end{array} \right) \right)^2.
\end{aligned}
$$

We have

$$
\begin{aligned}
\frac{1}{K} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} (W_{[h]} \hat{\gamma}_{\tilde{V}})^\top P_{[h,g]} \tilde{e}_{[g]} \right)^2 &= \frac{1}{K} \sum_{g \in [G]} \left((W_{[g]} \hat{\gamma}_{\tilde{V}})^\top P_{[g,g]} \tilde{e}_{[g]} \right)^2 \\
&\leqslant \frac{C \max_{1 \leqslant g \leqslant G} \left\| W_{[g]} \hat{\gamma}_{\tilde{V}} \right\|_2^2}{K} \sum_{g \in [G]} \left\| P_{[g,g]} \tilde{e}_{[g]} \right\|_2^2 \\
&= o_P(1),
\end{aligned}
$$

by Lemma B.1, where we use the fact that $W^\top P = 0$ and

$$\mathbb{E}\frac{1}{K}\sum_{g\in[G]}\left\|P_{[g,g]}\tilde{e}_{[g]}\right\|_2^2 \leqslant \frac{C}{K}\sum_{g\in[G]}\operatorname{trace}(P_{[g,g]}) = O(1).$$

Similarly, we have

$$\begin{aligned}
\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^\top P_{[h,g]}(W_{[g]}\hat{\gamma}_{\tilde{e}})\right)^2 &\leqslant \frac{\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2^2}{K}\sum_{g\in[G]}\left\|\sum_{h\in[G],h\neq g}P_{[g,h]}\tilde{V}_{[h]}\right\|_2^2 \\
&= o_P(1),
\end{aligned}$$

since

$$\begin{aligned}
\mathbb{E}\frac{1}{K}\sum_{g\in[G]}\left\|\sum_{h\in[G],h\neq g}P_{[g,h]}\tilde{V}_{[h]}\right\|_2^2 &= \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left\|P_{[g,h]}\tilde{V}_{[h]}\right\|_2^2 \\
&\leqslant \frac{C}{K}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}(P_{[g,h]}P_{[h,g]}) \\
&= O(1).
\end{aligned}$$

Finally, we have

$$\begin{aligned}
&\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}(W_{[h]}\hat{\gamma}_{\tilde{V}})^\top P_{[h,g]}(W_{[g]}\hat{\gamma}_{\tilde{e}})\right)^2 \\
&= \frac{1}{K}\sum_{g\in[G]}\left((W_{[g]}\hat{\gamma}_{\tilde{V}})^\top P_{[g,g]}(W_{[g]}\hat{\gamma}_{\tilde{e}})\right)^2 \\
&\leqslant \frac{C\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{V}}\right\|_2^2\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2^2}{K}\sum_{g\in[G]}\operatorname{trace}(P_{[g,g]}) \\
&= o_P(1).
\end{aligned}$$

Combining the above results with the triangle inequality, we have

$$\left(\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}V_{[h]}^{\top}P_{[h,g]}e_{[g]}\right)^2\right)^{1/2} = \left(\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2\right)^{1/2} + o_P(1).$$

In addition, we have

$$\mathbb{E}\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}P_{[h,g]}\tilde{e}_{[g]}\right)^2 = \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}P_{[h,g]}\tilde{e}_{[g]}\right)^2 \leqslant C||P||_F^2/K = O(1),$$

which implies

$$\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}P_{[h,g]}\tilde{e}_{[g]}\right)^2 = O_P(1).$$

Therefore, we have

$$\begin{aligned}
&\left|\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}V_{[h]}^{\top}P_{[h,g]}e_{[g]}\right)^2 - \frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2\right| \\
&= \left|\left(\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}V_{[h]}^{\top}P_{[h,g]}e_{[g]}\right)\right)^{1/2} - \left(\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2\right)^{1/2}\right| \\
&\times \left|\left(\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}V_{[h]}^{\top}P_{[h,g]}e_{[g]}\right)^2\right)^{1/2} + \left(\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2\right)^{1/2}\right| = o_P(1).
\end{aligned}$$

Next, we show that

$$\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2 - \frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2 = o_P(1). \tag{C.2}$$

By Markov inequality, it suffices to show that the RHS of the following display is $o(1)$

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{K}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^{2}-\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^{2}\right)\right)^{2}\\
&=\frac{1}{K^{2}}\mathbb{V}\left(\sum_{g,h,k\in[G]^{3},h\neq g,k\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\tilde{V}_{[k]}^{\top}P_{[k,g]}\tilde{e}_{[g]}\right)\\
&\leqslant\frac{C}{K^{2}}\mathbb{V}\left(\sum_{g,h\in[G]^{2},g\neq h}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^{2}\right)+\frac{C}{K^{2}}\mathbb{V}\left(\sum_{g,h,k\in[G]^{3},g\neq h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\tilde{V}_{[k]}^{\top}P_{[k,g]}\tilde{e}_{[g]}\right).
\end{aligned} \tag{C.3}$$

For the first term on the RHS of (C.3), we have

$$\begin{aligned}
&\frac{1}{K^{2}}\mathbb{V}\left(\sum_{g,h\in[G]^{2},g\neq h}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^{2}\right)\\
&=\frac{1}{K^{2}}\mathbb{E}\left(\sum_{g,h\in[G]^{2},g\neq h}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^{2}-\mathbb{E}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^{2}\right)^{2}\\
&\leqslant\frac{C}{K^{2}}\mathbb{E}\left(\sum_{g,h\in[G]^{2},g\neq h}\operatorname{trace}\left(P_{[g,h]}\left(\tilde{V}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_{h}^{\tilde{V},\tilde{V}}\right)P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_{g}^{\tilde{e},\tilde{e}}\right)\right)\right)^{2}\\
&+\frac{C}{K^{2}}\mathbb{E}\left(\sum_{g,h\in[G]^{2},g\neq h}\operatorname{trace}\left(P_{[g,h]}\Omega_{h}^{\tilde{V},\tilde{V}}P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_{g}^{\tilde{e},\tilde{e}}\right)\right)\right)^{2}\\
&+\frac{C}{K^{2}}\mathbb{E}\left(\sum_{g,h\in[G]^{2},g\neq h}\operatorname{trace}\left(\left(\tilde{V}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_{h}^{\tilde{V},\tilde{V}}\right)P_{[h,g]}\Omega_{g}^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\right)^{2}.
\end{aligned}$$

In addition, we have

$$\begin{aligned}
&\frac{1}{K^{2}}\mathbb{E}\left(\sum_{g,h\in[G]^{2},g\neq h}\operatorname{trace}\left(P_{[g,h]}\left(\tilde{V}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_{h}^{\tilde{V},\tilde{V}}\right)P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_{g}^{\tilde{e},\tilde{e}}\right)\right)\right)^{2}\\
&\leqslant\frac{C}{K^{2}}\sum_{g,h\in[G]^{2},g\neq h}\mathbb{E}\left(\operatorname{trace}\left(P_{[g,h]}\left(\tilde{V}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_{h}^{\tilde{V},\tilde{V}}\right)P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_{g}^{\tilde{e},\tilde{e}}\right)\right)\right)^{2}
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \frac{C}{K^2} \sum_{g,h\in[G]^2, g\neq h} \operatorname{trace}\left(P_{[g,h]}\mathbb{E}\left(\tilde{V}_{[h]}\tilde{V}_{[h]}^\top - \Omega_h^{\tilde{V},\tilde{V}}\right)^2 P_{[h,g]}\right) \\
&\qquad\qquad \times \operatorname{trace}\left(P_{[h,g]}\mathbb{E}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^\top - \Omega_g^{\tilde{e},\tilde{e}}\right)^2 P_{[g,h]}\right) \\
&\leqslant \frac{C}{K^2} \sum_{g,h\in[G]^2, g\neq h} \operatorname{trace}\left(P_{[g,h]}P_{[h,g]}\right) \\
&\leqslant \frac{C}{K^2}\operatorname{trace}\left((P-\bar{P})^2\right) \\
&= o(1),
\end{aligned}
$$

where the second inequality is by the trace Cauchy-Schwartz inequality (e.g., Magnus and Neudecker (2019)). Similarly, we have

$$
\begin{aligned}
&\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2, g\neq h} \operatorname{trace}\left(P_{[g,h]}\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^\top - \Omega_g^{\tilde{e},\tilde{e}}\right)\right)\right)^2 \\
&= \frac{1}{K^2}\mathbb{E}\left(\sum_{g\in[G]} \operatorname{trace}\left(\left(\sum_{h\neq g} P_{[g,h]}\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\right)\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^\top - \Omega_g^{\tilde{e},\tilde{e}}\right)\right)\right)^2 \\
&= \frac{1}{K^2}\sum_{g\in[G]} \mathbb{E}\left(\operatorname{trace}\left(\left(\sum_{h\neq g} P_{[g,h]}\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\right)\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^\top - \Omega_g^{\tilde{e},\tilde{e}}\right)\right)\right)^2 \\
&\leqslant \frac{1}{K^2}\sum_{g\in[G]} \operatorname{trace}\left(\left(\sum_{h\neq g} P_{[g,h]}\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\right)^2\right)\operatorname{trace}\left(\mathbb{E}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^\top - \Omega_g^{\tilde{e},\tilde{e}}\right)^2\right) \\
&\leqslant \frac{C}{K^2}\sum_{g\in[G]} \operatorname{trace}\left(\sum_{h\neq g} P_{[g,h]}\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\right) \\
&\leqslant \frac{C}{K^2} \sum_{g,h\in[G]^2, g\neq h} \operatorname{trace}\left(P_{[g,h]}P_{[h,g]}\right) \\
&= o(1),
\end{aligned}
$$

where the first inequality is by the trace Cauchy-Schwartz inequality, and, following the same argument,

$$\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\text{trace}\left(\left(\tilde{V}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_h^{\tilde{V},\tilde{V}}\right)P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\right)^2=o(1).$$

Combining these bounds with (C.3), we have

$$\frac{1}{K^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2\right)=o(1). \tag{C.4}$$

Now consider the second term on the RHS of (C.3). We have

$$\begin{aligned}
&\frac{1}{K^2}\mathbb{V}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\tilde{V}_{[k]}^{\top}P_{[k,g]}\tilde{e}_{[g]}\right)\\
&=\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\tilde{V}_{[k]}^{\top}P_{[k,g]}\tilde{e}_{[g]}\right)^2\\
&\leqslant\frac{C}{K^2}\mathbb{E}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{e}}\right)P_{[g,k]}\tilde{V}_{[k]}\right)^2\\
&+\frac{C}{K^2}\mathbb{E}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\tilde{V}_{[k]}\right)^2,
\end{aligned}$$

where

$$\begin{aligned}
&\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{e}}\right)P_{[g,k]}\tilde{V}_{[k]}\right)^2\\
&\leqslant\frac{C}{K^2}\sum_{g,h,k\in[G]^3,g\neq h\neq k}\mathbb{E}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{e}}\right)P_{[g,k]}\tilde{V}_{[k]}\right)^2\\
&\leqslant\frac{C}{K^2}\sum_{g,h\in[G]^2,g\neq h}\text{trace}\left(P_{[g,h]}P_{[h,g]}\right)\\
&=o(1),
\end{aligned}$$

and

$$
\begin{aligned}
&\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^\top P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\tilde{V}_{[k]}\right)^2\\
&=\frac{1}{K^2}\mathbb{E}\left(\sum_{h,k\in[G]^2,h\neq k}\tilde{V}_{[h]}^\top\left(\sum_{g\neq h\neq k}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\right)\tilde{V}_{[k]}\right)^2\\
&\leqslant\frac{C}{K^2}\sum_{h,k\in[G]^2,h\neq k}\mathbb{E}\left(\tilde{V}_{[h]}^\top\left(\sum_{g\neq h\neq k}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\right)\tilde{V}_{[k]}\right)^2\\
&\leqslant\frac{C}{K^2}\sum_{h,k\in[G]^2,h\neq k}\operatorname{trace}\left(\left((P-\bar{P})\Omega_{\tilde{e}}(P-\bar{P})\right)_{[h,k]}\left((P-\bar{P})\Omega_{\tilde{e}}(P-\bar{P})\right)_{[k,h]}\right)\\
&\leqslant\frac{C}{K^2}\operatorname{trace}\left((P-\bar{P})\Omega_{\tilde{e}}(P-\bar{P})(P-\bar{P})\Omega_{\tilde{e}}(P-\bar{P})\right)\\
&=o(1).
\end{aligned}
$$

This implies that

$$
\frac{1}{K^2}\mathbb{V}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\tilde{V}_{[k]}^\top P_{[k,g]}\tilde{e}_{[g]}\right)=o(1). \tag{C.5}
$$

Combining (C.3)–(C.5), we have established (C.2), which further implies the desired result that

$$
\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}V_{[h]}^\top P_{[h,g]}e_{[g]}\right)^2=\frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1).
$$

Note that, by Lemma B.2 and the fact that $W^\top Q=0$, we can show

$$
\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}V_{[h]}^\top Q_{[h,g]}e_{[g]}\right)^2=\frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1)
$$

in the same manner by replacing P by Q.

Next, we note that

$$
\begin{aligned}
&\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(V_{[h]}^\top P_{[h,g]}e_{[g]}\right)\left(V_{[g]}^\top P_{[g,h]}e_{[h]}\right)\\
&=\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(\underbrace{\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}}_{U_{gh}^{(1)}}-\underbrace{(W_{[h]}\hat{\gamma}_{\tilde{V}})^\top P_{[h,g]}\tilde{e}_{[g]}}_{U_{gh}^{(2)}}-\underbrace{\tilde{V}_{[h]}^\top P_{[h,g]}(W_{[g]}\hat{\gamma}_{\tilde{e}})}_{U_{gh}^{(3)}}+\underbrace{(W_{[h]}\hat{\gamma}_{\tilde{V}})^\top P_{[h,g]}(W_{[g]}\hat{\gamma}_{\tilde{e}})}_{U_{gh}^{(4)}}\right)\\
&\qquad\times\left(\underbrace{\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}}_{U_{hg}^{(1)}}-\underbrace{(W_{[g]}\hat{\gamma}_{\tilde{V}})^\top P_{[g,h]}\tilde{e}_{[h]}}_{U_{hg}^{(2)}}-\underbrace{\tilde{V}_{[g]}^\top P_{[g,h]}(W_{[h]}\hat{\gamma}_{\tilde{e}})}_{U_{hg}^{(3)}}+\underbrace{(W_{[g]}\hat{\gamma}_{\tilde{V}})^\top P_{[g,h]}(W_{[h]}\hat{\gamma}_{\tilde{e}})}_{U_{hg}^{(4)}}\right),
\end{aligned}
$$

and by using a similar argument as in the proof for (C.1), we have

$$
\begin{aligned}
\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{gh}^{(1),2}&=\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{hg}^{(1),2}=O_P(1),\\
\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{gh}^{(2),2}&=\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{hg}^{(2),2}=o_P(1),\\
\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{gh}^{(3),2}&=\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{hg}^{(3),2}=o_P(1),\\
\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{gh}^{(4),2}&=\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{hg}^{(4),2}=o_P(1).
\end{aligned}
$$

Then by repeatedly applying Cauchy-Schwarz inequality, we have

$$
\left|\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{gh}^{(s_1)}U_{hg}^{(s_2)}\right|\leqslant\left(\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{gh}^{(s_1),2}\right)^{1/2}\left(\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}U_{gh}^{(s_2),2}\right)^{1/2}=o_P(1)
$$

for $s_1\neq 1$ or $s_2\neq 1$, whence

$$
\begin{aligned}
&\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(V_{[h]}^\top P_{[h,g]}e_{[g]}\right)\left(V_{[g]}^\top P_{[g,h]}e_{[h]}\right)\\
&=\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)\left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)+o_P(1).
\end{aligned}
$$

Next, we note that

$$
\begin{aligned}
&\mathbb{V}\left(\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\right)\\
&=\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\right.\\
&\qquad\qquad\left.-\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\right)^2\\
&\leqslant\frac{C}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(P_{[g,h]}\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_h^{\tilde{e},\tilde{V}}\right)P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\right)\right)^2\\
&+\frac{C}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(P_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\right)\right)^2\\
&+\frac{C}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_h^{\tilde{e},\tilde{V}}\right)P_{[h,g]}\Omega_g^{\tilde{e},\tilde{V}}P_{[g,h]}\right)\right)^2\\
&=o(1),
\end{aligned}
$$

where the last equality holds because

$$
\begin{aligned}
&\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(P_{[g,h]}\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_h^{\tilde{e},\tilde{V}}\right)P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\right)\right)^2\\
&\leqslant\frac{C}{K^2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\operatorname{trace}\left(P_{[g,h]}\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_h^{\tilde{e},\tilde{V}}\right)P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\right)\right)^2\\
&\leqslant\frac{C}{K^2}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(P_{[g,h]}\mathbb{E}\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_h^{\tilde{e},\tilde{V}}\right)\left(\tilde{V}_{[h]}\tilde{e}_{[h]}^{\top}-\Omega_h^{\tilde{V},\tilde{e}}\right)P_{[h,g]}\right)\\
&\qquad\qquad\times\operatorname{trace}\left(P_{[h,g]}\mathbb{E}\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\left(\tilde{V}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_g^{\tilde{V},\tilde{e}}\right)P_{[g,h]}\right)\\
&\leqslant\frac{C}{K^2}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(P_{[g,h]}P_{[h,g]}\right)\\
&=o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(P_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}P_{[h,g]}\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\right)\right)^2\\
&=\frac{1}{K^2}\mathbb{E}\left(\sum_{g\in[G]}\operatorname{trace}\left(\left(\sum_{h\neq g}P_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}P_{[h,g]}\right)\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\right)\right)^2\\
&=\frac{1}{K^2}\sum_{g\in[G]}\mathbb{E}\left(\operatorname{trace}\left(\left(\sum_{h\neq g}P_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}P_{[h,g]}\right)\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\right)\right)^2\\
&\leqslant\frac{1}{K^2}\sum_{g\in[G]}\operatorname{trace}\left(\left(\sum_{h\neq g}P_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}P_{[h,g]}\right)\left(\sum_{k\neq g}P_{[g,k]}\Omega_k^{\tilde{V},\tilde{e}}P_{[k,g]}\right)\right)\\
&\qquad\qquad\times\operatorname{trace}\left(\mathbb{E}\left(\tilde{e}_{[g]}\tilde{V}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{V}}\right)\left(\tilde{V}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_g^{\tilde{V},\tilde{e}}\right)\right)\\
&\leqslant\frac{C}{K^2}\sum_{g,h,k\in[G]^2,h\neq g,k\neq g}\operatorname{trace}\left(P_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}P_{[h,g]}P_{[g,k]}\Omega_k^{\tilde{V},\tilde{e}}P_{[k,g]}\right)\\
&\leqslant\frac{C}{K^2}\sum_{g,h,k\in[G]^2,h\neq g,k\neq g}\left(\operatorname{trace}\left(P_{[k,g]}P_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}\Omega_h^{\tilde{V},\tilde{e}}P_{[h,g]}P_{[g,k]}\right)\right)^{1/2}\\
&\qquad\qquad\qquad\times\left(\operatorname{trace}\left(P_{[h,g]}P_{[g,k]}\Omega_k^{\tilde{V},\tilde{e}}\Omega_k^{\tilde{e},\tilde{V}}P_{[k,g]}P_{[g,h]}\right)\right)^{1/2}\\
&\leqslant\frac{C}{K^2}\sum_{g,h,k\in[G]^2,h\neq g,k\neq g}\operatorname{trace}\left(P_{[k,g]}P_{[g,h]}P_{[h,g]}P_{[g,k]}\right)\\
&\leqslant\frac{C}{K^2}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(P_{[g,h]}P_{[h,g]}\right)\\
&=o(1),
\end{aligned}
$$

and, following the same argument,

$$
\frac{1}{K^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^{\top}-\Omega_h^{\tilde{e},\tilde{V}}\right)P_{[h,g]}\Omega_g^{\tilde{e},\tilde{V}}P_{[g,h]}\right)\right)^2=o(1).
$$

This implies that

$$
\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)
$$

$$= \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[h]}^\top P_{[h,g]} \tilde{e}_{[g]} \right) \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]} \right) + o_P(1).$$

Note also that we can replace P by Q, as in the proof for (C.1) and (C.2), and this concludes the proof for the first result.

For the second result, we focus on the case when $\tilde{u} = \tilde{V}$ and $\tilde{v} = \tilde{e}$. Recall that

$$Q = M_W(P - \bar{P})M_W = P - \bar{P} + P_W \bar{P} + \bar{P} P_W - P_W \bar{P} P_W,$$

which implies that

$$Q_{[h,g]} = P_{[h,g]} + P_{W,[h,g]} P_{[g,g]} + P_{[h,h]} P_{W,[h,g]} - \sum_{k \in [G]} P_{W,[h,k]} P_{[k,k]} P_{W,[k,g]}, \quad g \neq h.$$

Therefore, we have

$$\begin{aligned}
&\frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 \\
&= \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[h]}^\top \left(P_{[h,g]} + P_{W,[h,g]} P_{[g,g]} + P_{[h,h]} P_{W,[h,g]} - \sum_{k \in [G]} P_{W,[h,k]} P_{[k,k]} P_{W,[k,g]} \right) \tilde{e}_{[g]} \right)^2,
\end{aligned}$$

where

$$\frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[h]}^\top P_{W,[h,g]} P_{[g,g]} \tilde{e}_{[g]} \right)^2 \leqslant \frac{C}{K} \sum_{g,h \in [G]^2, g \neq h} \operatorname{trace} \left(P_{W,[g,h]} P_{W,[h,g]} \right) = o(1),$$

since d_w is fixed,

$$\frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top P_{[g,g]} P_{W,[g,h]} \tilde{e}_{[g]} \right)^2 = o(1)$$

by the same argument as above, and

$$
\begin{aligned}
&\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^{\top}\left(\sum_{k\in[G]}P_{W,[h,k]}P_{[k,k]}P_{W,[k,g]}\right)\tilde{e}_{[g]}\right)^2\\
&\leqslant\frac{C}{K}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(\left(P_W\bar{P}P_W\right)_{[g,h]}\left(P_W\bar{P}P_W\right)_{[h,g]}\right)\\
&\leqslant\frac{C}{K}\operatorname{trace}\left(P_W\bar{P}P_WP_W\bar{P}P_W\right)\\
&=o(1).
\end{aligned}
$$

It follows that

$$
\frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2=\frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2+o(1).
$$

Similarly, we can show that

$$
\begin{aligned}
&\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\\
&=\frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)\left(\tilde{V}_{[g]}^{\top}Q_{[g,h]}\tilde{e}_{[h]}\right)+o(1).
\end{aligned}
$$

This concludes the proof. □

C.5 Proof of Lemma B.5

We prove each result in turn. To begin with, we note that

$$
\Sigma=\mathbb{V}\left(\sum_{g\in[G]}\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)+\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right),
$$

where

$$\mathbb{V}\left(\sum_{g\in[G]}\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)=\mathbb{E}\left(\hat{\Pi}^{\top}\tilde{e}\right)\geqslant\frac{1}{C}\hat{\Pi}^{\top}\hat{\Pi}\geqslant\frac{1}{C}\Pi^{\top}\Pi,$$

and

$$\begin{aligned}
&\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)^2\\
&=\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)^2+\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)\\
&=\frac{1}{2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left[\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)+\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\right]^2\\
&=\frac{1}{2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left[\begin{pmatrix}\tilde{e}_{[g]} & \tilde{V}_{[g]}\end{pmatrix}\begin{pmatrix}0_{n_g\times n_h} & P_{[g,h]}\\ P_{[g,h]} & 0_{n_g\times n_h}\end{pmatrix}\begin{pmatrix}\tilde{e}_{[h]}\\ \tilde{V}_{[h]}\end{pmatrix}\right]^2\\
&\geqslant\frac{1}{2C}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left[\begin{pmatrix}0_{n_g\times n_h} & P_{[g,h]}\\ P_{[g,h]} & 0_{n_g\times n_h}\end{pmatrix}\begin{pmatrix}0_{n_h\times n_g} & P_{[h,g]}\\ P_{[h,g]} & 0_{n_h\times n_g}\end{pmatrix}\right]\\
&=\frac{1}{C}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left[P_{[g,h]}P_{[h,g]}\right]\\
&=\frac{1}{C}\sum_{g\in[G]}\operatorname{trace}\left[P_{[g,g]}-P_{[g,g]}^2\right]\\
&\geqslant\frac{1}{C}\sum_{g\in[G]}\left(1-\lambda_{\max}(P_{[g,g]})\right)\operatorname{trace}\left[P_{[g,g]}\right]\\
&\geqslant\frac{1}{C}K.
\end{aligned}$$

These two lower bounds imply the desired result that

$$\Sigma\geqslant(\Pi^{\top}\Pi+K)/C.$$

Next, we note that

$$
\begin{aligned}
&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^{2}\\
=&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^{2}\\
+&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^{2}\\
+&\frac{2}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{\Pi}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right).
\end{aligned}
\tag{C.6}
$$

For the first term on the RHS of (C.6), we have

$$
\mathbb{E}\frac{1}{\Sigma}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^{2}-\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)\right)=0,
$$

and by Assumption 3

$$
\begin{aligned}
&\mathbb{V}\left(\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^{2}\right)\\
=&\frac{1}{\Sigma^{2}}\mathbb{V}\left(\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)^{2}\right)\\
\leqslant&\frac{1}{\Sigma^{2}}\sum_{g\in[G]}\mathbb{E}\left(\bar{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)^{4}\\
\leqslant&\frac{C}{\Sigma^{2}}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\bar{\Pi}_{[g]}\right)^{2}\\
\leqslant&\frac{C\max_{1\leqslant g\leqslant G}\left\|\bar{\Pi}_{[g]}\right\|_{2}^{2}\bar{\Pi}^{\top}\bar{\Pi}}{\left(\Pi^{\top}\Pi+K\right)^{2}}\\
=&o(1).
\end{aligned}
$$

Therefore, we have

$$\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2=\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1). \tag{C.7}$$

For the second term on the RHS of (C.6), by Lemma B.4 and the fact that $K/\Sigma=O(1)$, we have

$$\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2=\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1). \tag{C.8}$$

For the last term on the RHS of (C.6), we have

$$\mathbb{E}\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{k\in[G],k\neq g}\tilde{\Pi}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right)\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)=0,$$

and

$$\begin{aligned}
&\mathbb{V}\left(\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{k\in[G],k\neq g}\tilde{\Pi}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right)\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)\right)\\
&=\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2\\
&\leqslant\frac{C}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{e}}\right)Q_{[g,h]}\tilde{V}_{[h]}\right)^2\\
&+\frac{C}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}\Omega_g^{\tilde{e},\tilde{e}}Q_{[g,h]}\tilde{V}_{[h]}\right)^2,
\end{aligned}$$

where

$$\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}\left(\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}-\Omega_g^{\tilde{e},\tilde{e}}\right)Q_{[g,h]}\tilde{V}_{[h]}\right)^2$$

$$
\begin{aligned}
&\leqslant \frac{C}{\Sigma^2} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\bar{\Pi}_{[g]}^\top \left(\tilde{e}_{[g]} \tilde{e}_{[g]}^\top - \Omega_g^{\tilde{e},\tilde{e}} \right) Q_{[g,h]} \tilde{V}_{[h]} \right)^2 \\
&\leqslant \frac{C \max_{1 \leqslant g \leqslant G} \left\| \bar{\Pi}_{[g]} \right\|_2^2}{\Sigma^2} \sum_{g,h \in [G]^2, g \neq h} \operatorname{trace} \left(Q_{[g,h]} Q_{[h,g]} \right) \\
&= o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{\Sigma^2} \mathbb{E} \left(\sum_{g,h \in [G]^2, g \neq h} \bar{\Pi}_{[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} Q_{[g,h]} \tilde{V}_{[h]} \right)^2 \\
&= \frac{1}{\Sigma^2} \mathbb{E} \left(\bar{\Pi}^\top \Omega_{\tilde{e}} (Q - \bar{Q}) \tilde{V} \right)^2 \\
&\leqslant \frac{C \bar{\Pi}^\top \bar{\Pi}}{\left(\Pi^\top \Pi + K \right)^2} \\
&= o(1).
\end{aligned}
$$

Therefore, we have

$$
\frac{2}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{\Pi}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right) = o_P(1). \tag{C.9}
$$

Combining (C.6)–(C.9), we have the desired result that

$$
\begin{aligned}
&\frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 \\
&= \frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 + \frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 + o_P(1) \\
&= \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 + \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 + o_P(1) \\
&= \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 + o_P(1).
\end{aligned}
$$

Next, we note that

$$
\begin{aligned}
&\frac{1}{\Sigma}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right)^2-\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2\right)\\
&=\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)^2 &\text{(C.10)}\\
&-\frac{2}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right). &\text{(C.11)}
\end{aligned}
$$

For the first term on the RHS of (C.10), we have

$$
\begin{aligned}
&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)^2\\
&\leqslant \max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2^2\times\frac{1}{\Sigma}\sum_{g\in[G]}\left\|\sum_{h\in[G],h\neq g}Q_{[g,h]}\tilde{X}_{[h]}\right\|_2^2\\
&\leqslant \max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2^2\times\frac{C}{\Sigma}\sum_{g\in[G]}\left(\left\|\bar{\Pi}_{[g]}\right\|_2^2+\left\|\sum_{h\in[G],h\neq g}Q_{[g,h]}\tilde{V}_{[h]}\right\|_2^2\right)\\
&=o_P(1),
\end{aligned}
$$

by Lemma B.1. For the second term on the RHS of (C.10), we have

$$
\begin{aligned}
&\left|\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right)\right|\\
&\leqslant\left(\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)^2\right)^{1/2}\times\left(\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right)^2\right)^{1/2}\\
&=o_P(1).
\end{aligned}
$$

Combining the two bounds above, we have the desired result that

$$\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right)^2=\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1).$$

Next, we note that

$$\begin{aligned}
&\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right)^2\\
&=\underbrace{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}X_{[g]}\right)^2}_{R_1}\\
&+\underbrace{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right)^2}_{R_2}\\
&+2\times\underbrace{\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}X_{[g]}\tilde{V}_{[h]}^{\top}Q_{[h,g]}X_{[g]}}_{R_3}.
\end{aligned}\tag{C.12}$$

For R_1, we have

$$\begin{aligned}
R_1&=\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)^2+\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}V_{[g]}\right)^2+\frac{2}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)\left(\bar{\Pi}_{[g]}^{\top}V_{[g]}\right)\\
&=\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)^2+\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\tilde{V}_{[g]}\right)^2+\frac{2}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)\left(\bar{\Pi}_{[g]}^{\top}\tilde{V}_{[g]}\right)+o_P(1)\\
&=\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)^2+\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\bar{\Pi}_{[g]}^{\top}\tilde{V}_{[g]}\right)^2+o_P(1),
\end{aligned}$$

where the second equality holds by using $V_{[g]}=\tilde{V}_{[g]}-W_{[g]}\hat{\gamma}_{\tilde{V}}$ and Lemma B.1, and the last equality holds because

$$\mathbb{V}\left(\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\tilde{V}_{[g]}\right)^2\right)$$

$$
\begin{aligned}
&\leqslant \frac{1}{\Sigma^2} \sum_{g \in [G]} \mathbb{E}\left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\right)^4 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g \in [G]} \left(\bar{\Pi}_{[g]}^\top \bar{\Pi}_{[g]}\right)^2 \\
&\leqslant \frac{C \max_{g \in [G]} \left\|\bar{\Pi}_{[g]}\right\|_2^2 \bar{\Pi}^\top \bar{\Pi}}{(\Pi^\top \Pi + K)^2} \\
&= o(1)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{V}\left(\frac{1}{\Sigma} \sum_{g \in [G]} \left(\bar{\Pi}_{[g]}^\top \Pi_{[g]}\right) \left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\right)\right) \\
&= \frac{1}{\Sigma^2} \sum_{g \in [G]} \mathbb{E}\left(\left(\bar{\Pi}_{[g]}^\top \Pi_{[g]}\right) \left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\right)\right)^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g \in [G]} \left(\bar{\Pi}_{[g]}^\top \Pi_{[g]}\right)^2 \left(\bar{\Pi}_{[g]}^\top \bar{\Pi}_{[g]}\right) \\
&\leqslant \frac{C \max_{g \in [G]} \left\|\bar{\Pi}_{[g]}\right\|_2^2 \bar{\Pi}^\top \bar{\Pi}}{(\Pi^\top \Pi + K)^2} \\
&= o(1).
\end{aligned}
$$

For R_2, we have

$$
\begin{aligned}
R_2 = &\underbrace{\frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \Pi_{[g]}\right)^2}_{R_{2,1}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} V_{[g]}\right)^2}_{R_{2,2}} \\
&+ 2 \times \underbrace{\frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \Pi_{[g]}\right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^\top Q_{[k,g]} V_{[g]}\right)}_{R_{2,3}}.
\end{aligned}
$$

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

$$\mathbb{V}\left(R_{2,1}\right) \leqslant \frac{C}{\Sigma^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)^2\right) + \frac{C}{\Sigma^2}\mathbb{V}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)\left(\tilde{V}_{[k]}^\top Q_{[k,g]}\Pi_{[g]}\right)\right),$$

where

$$\begin{aligned}
&\frac{1}{\Sigma^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)^2\right)\\
&=\frac{1}{\Sigma^2}\mathbb{V}\left(\sum_{h\in[G]}\tilde{V}_{[h]}^\top\left(\sum_{g\in[G],g\neq h}Q_{[h,g]}\Pi_{[g]}\Pi_{[g]}^\top Q_{[g,h]}\right)\tilde{V}_{[h]}\right)^2\\
&\leqslant\frac{C}{\Sigma^2}\sum_{h\in[G]}\left\|\sum_{g\in[G],g\neq h}Q_{[h,g]}\Pi_{[g]}\Pi_{[g]}^\top Q_{[g,h]}\right\|_{op}^2\\
&\leqslant\frac{C}{\Sigma^2}\sum_{h\in[G]}\text{trace}\left(\sum_{g\in[G],g\neq h}Q_{[h,g]}\Pi_{[g]}\Pi_{[g]}^\top Q_{[g,h]}\right)\\
&\leqslant\frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\text{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)\\
&=o(1),
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{\Sigma^2}\mathbb{V}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)\left(\tilde{V}_{[k]}^\top Q_{[k,g]}\Pi_{[g]}\right)\right)\\
&=\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)\left(\tilde{V}_{[k]}^\top Q_{[k,g]}\Pi_{[g]}\right)\right)^2\\
&\leqslant\frac{C}{\Sigma^2}\sum_{g,h,k\in[G]^3,g\neq h\neq k}\mathbb{E}\left(\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)\left(\tilde{V}_{[k]}^\top Q_{[k,g]}\Pi_{[g]}\right)\right)^2
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \frac{C}{\Sigma^2} \sum_{g,h,k \in [G]^3, g \neq h \neq k} \operatorname{trace}\left(Q_{[h,g]} Q_{[g,h]}\right) tr\left(Q_{[k,g]} Q_{[g,k]}\right) \\
&= o(1).
\end{aligned}
$$

This implies that

$$
R_{2,1} = \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^{\top} Q_{[h,g]} \Pi_{[g]} \right)^2 + o_P(1).
$$

For $R_{2,2}$, since $K/\Sigma = O(1)$, by Lemma B.6, we have

$$
R_{2,2} = \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]} \right)^2 + o_P(1).
$$

For $R_{2,3}$, we have

$$
\begin{aligned}
R_{2,3} &= \frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^{\top} Q_{[h,g]} \Pi_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^{\top} Q_{[k,g]} \tilde{V}_{[g]} \right) + o_P(1) \\
&= o_P(1),
\end{aligned}
$$

where the second equality holds because

$$
\mathbb{E} \frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^{\top} Q_{[h,g]} \Pi_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^{\top} Q_{[k,g]} \tilde{V}_{[g]} \right) = 0,
$$

$$
\begin{aligned}
&\mathbb{V} \left(\frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^{\top} Q_{[h,g]} \Pi_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^{\top} Q_{[k,g]} \tilde{V}_{[g]} \right) \right) \\
&\leqslant \frac{C}{\Sigma^2} \mathbb{V} \left(\sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[h]}^{\top} Q_{[h,g]} \Pi_{[g]} \tilde{V}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]} \right)
\end{aligned}
$$

$$+\frac{C}{\Sigma^2}\mathbb{V}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\tilde{V}_{[k]}^\top Q_{[k,g]}\tilde{V}_{[g]}\right),$$

where

$$\begin{aligned}
&\frac{1}{\Sigma^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)\\
&\leqslant \frac{C}{\Sigma^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\Pi_{[g]}^\top Q_{[g,h]}\Omega_h^{\tilde{V},\tilde{V}}Q_{[h,g]}\tilde{V}_{[g]}\right)^2\\
&+\frac{C}{\Sigma^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\Pi_{[g]}^\top Q_{[g,h]}(\tilde{V}_{[h]}\tilde{V}_{[h]}^\top-\Omega_h^{\tilde{V},\tilde{V}})Q_{[h,g]}\tilde{V}_{[g]}\right)^2\\
&\leqslant \frac{C}{\Sigma^2}\sum_{g\in[G]}\left\|\left(\sum_{h\in[G],h\neq g}Q_{[g,h]}\Omega_h^{\tilde{V},\tilde{V}}Q_{[h,g]}\right)\Pi_{[g]}\right\|_2^2\\
&+\frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)^2\\
&\leqslant \frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\text{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)\\
&=o(1),
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{\Sigma^2}\mathbb{V}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\tilde{V}_{[k]}^\top Q_{[k,g]}\tilde{V}_{[g]}\right)\\
&=\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h,k\in[G]^3,g\neq h\neq k}\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\tilde{V}_{[k]}^\top Q_{[k,g]}\tilde{V}_{[g]}\right)^2\\
&\leqslant \frac{C}{\Sigma^2}\sum_{g,h,k\in[G]^3,g\neq h\neq k}\mathbb{E}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\tilde{V}_{[k]}^\top Q_{[k,g]}\tilde{V}_{[g]}\right)^2\\
&\leqslant \frac{C}{\Sigma^2}\sum_{g,h,k\in[G]^3,g\neq h\neq k}\text{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)\text{trace}\left(Q_{[k,g]}Q_{[g,k]}\right)\\
&=o(1).
\end{aligned}$$

Combining the results above, we have

$$R_2 = \frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)^2 + \frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{V}_{[g]}\right)^2 + o_P(1).$$

For R_3, we have

$$\begin{aligned}
R_3 = & \underbrace{\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}}_{R_{3,1}} \\
& + \underbrace{\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\tilde{V}_{[h]}^{\top}Q_{[h,g]}V_{[g]}}_{R_{3,2}} \\
& + \underbrace{\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}V_{[g]}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}}_{R_{3,3}} \\
& + \underbrace{\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}V_{[g]}\tilde{V}_{[h]}^{\top}Q_{[h,g]}V_{[g]}}_{R_{3,4}}.
\end{aligned}$$

For $R_{3,1}$, it has mean zero and

$$\begin{aligned}
\mathbb{V}\left(R_{3,1}\right) &= \frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\left\|\sum_{g\in[G],g\neq h}Q_{[h,g]}\Pi_{[g]}\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right\|_2^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\left\|\sum_{g\in[G]}Q_{[h,g]}\Pi_{[g]}\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right\|_2^2 + \frac{C}{\Sigma^2}\sum_{h\in[G]}\left\|Q_{[h,h]}\Pi_{[h]}\bar{\Pi}_{[h]}^{\top}\Pi_{[h]}\right\|_2^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\left\|\sum_{g\in[G]}Q_{[h,g]}\Pi_{[g]}\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right\|_2^2 + o(1) \\
&= o(1),
\end{aligned}$$

where the last equality holds because

$$
\begin{aligned}
&\frac{1}{\Sigma^2}\sum_{h\in[G]}\left\|\sum_{g\in[G]} Q_{[h,g]}\Pi_{[g]}\bar{\Pi}_{[g]}^\top\Pi_{[g]}\right\|_2^2\\
&=\frac{1}{\Sigma^2}\sum_{g,g'\in[G]^2}(\Pi_{[g]}\bar{\Pi}_{[g]}^\top\Pi_{[g]})^\top\left(\sum_{h\in[G]} Q_{[g,h]}Q_{[h,g']}\right)\Pi_{[g']}\bar{\Pi}_{[g']}^\top\Pi_{[g']}\\
&\leqslant\frac{1}{\Sigma^2}\sum_{g\in[G]}\left\|\Pi_{[g]}\bar{\Pi}_{[g]}^\top\Pi_{[g]}\right\|_2^2\\
&=o(1).
\end{aligned}
$$

For $R_{3,2}$, we have

$$
R_{3,2}=\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^\top\Pi_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}+o_P(1),
$$

where the first term has mean zero and

$$
\begin{aligned}
&\mathbb{V}\left(\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^\top\Pi_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)\\
&=\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^\top\Pi_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)^2\\
&\leqslant\frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\bar{\Pi}_{[g]}^\top\Pi_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)^2\\
&\leqslant\frac{C\max_{g\in[G]}\left\|\bar{\Pi}_{[g]}\right\|_2^2}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)\\
&=o(1).
\end{aligned}
$$

By using the same argument, we also have $R_{3,3}=o_P(1)$. For $R_{3,4}$, we have

$$
R_{3,4}=\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^\top\tilde{V}_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}+o_P(1),
$$

where the first term has mean zero and

$$
\begin{aligned}
&\mathbb{V}\left(\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)\\
&\leqslant \frac{C}{\Sigma^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^\top \Omega_g^{\tilde{V},\tilde{V}} Q_{[g,h]}\tilde{V}_{[h]}\right)\\
&+\frac{C}{\Sigma^2}\mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\bar{\Pi}_{[g]}^\top (\tilde{V}_{[g]}\tilde{V}_{[g]}^\top - \Omega_g^{\tilde{V},\tilde{V}}) Q_{[g,h]}\tilde{V}_{[h]}\right)\\
&\leqslant \frac{C}{\Sigma^2}\mathbb{E}\left(\bar{\Pi}^\top \Omega_{\tilde{V}}(Q-\bar{Q})\tilde{V}\right)^2\\
&+\frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)^2\\
&\leqslant \frac{C\bar{\Pi}^\top\bar{\Pi}}{\Sigma^2}+\frac{C\max_{g\in[G]}\left\|\bar{\Pi}_{[g]}\right\|_2^2}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)\\
&= o(1).
\end{aligned}
$$

Combining the results above, we have $R_3 = o_P(1)$.

Therefore, by combining (C.12) with the calculations about terms R_1 to R_3, we have

$$
\begin{aligned}
\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}X_{[g]}\right)^2 &= \frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^\top \Pi_{[g]}\right)^2 + \frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\bar{\Pi}_{[g]}^\top\tilde{V}_{[g]}\right)^2\\
&+\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)^2\\
&+\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)^2 + o_P(1).
\end{aligned}
$$

Finally, we note that

$$
\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}X_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^\top Q_{[k,g]}e_{[g]}\right)
$$

$$
\begin{aligned}
&= \underbrace{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}e_{[g]}\right)}_{R_4} \\
&+ \underbrace{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}V_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}e_{[g]}\right)}_{R_5}. \qquad \text{(C.13)}
\end{aligned}
$$

For R_4, we have

$$
\begin{aligned}
R_4 &= \frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right) + o_P(1) \\
&= \underbrace{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)\left(\bar{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)}_{R_{4,1}} \\
&+ \underbrace{\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)\left(\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)}_{R_{4,2}} \\
&+ \underbrace{\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{V}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)\left(\bar{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)}_{R_{4,3}} \\
&+ \underbrace{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{V}_{[k]}^{\top}Q_{[k,g]}\tilde{e}_{[g]}\right)}_{R_{4,4}} + o_P(1).
\end{aligned}
$$

By using a similar argument as in the proof for R_1, we have $R_{4,1} = o_P(1)$. For $R_{4,2}$, it has mean zero and

$$
\begin{aligned}
\mathbb{V}\left(R_{4,2}\right) &= \frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)\left(\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\left(\bar{\Pi}_{[g]}^{\top}\Pi_{[g]}\right)^2\mathbb{E}\left(\tilde{V}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2
\end{aligned}
$$

$$\leqslant \frac{C \max_{g\in[G]} \left\|\bar{\Pi}_{[g]}\right\|_2^2}{\Sigma^2} \sum_{g,h\in[G]^2, g\neq h} \operatorname{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)$$

$$= o(1).$$

Therefore, we have $R_{4,2} = o_P(1)$. Using the same argument, we also have $R_{4,3} = o_P(1)$. In addition, by using a similar argument as in the proof for $R_{2,3}$, we have $R_{4,4} = o_P(1)$, which implies $R_4 = o_P(1)$.

For R_5, we have

$$\begin{aligned}
R_5 &= \frac{1}{\Sigma} \sum_{g\in[G]} \left(\sum_{h\in[G], h\neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \left(\sum_{k\in[G], k\neq g} \tilde{X}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right) + o_P(1) \\
&= \underbrace{\frac{1}{\Sigma} \sum_{g\in[G]} \left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\right) \left(\bar{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)}_{R_{5,1}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h\in[G]^2, g\neq h} \left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]}\right)}_{R_{5,2}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h\in[G]^2, g\neq h} \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]}\right) \left(\bar{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)}_{R_{5,3}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g\in[G]} \left(\sum_{h\in[G], h\neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \left(\sum_{k\in[G], k\neq g} \tilde{V}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right)}_{R_{5,4}} + o_P(1).
\end{aligned}$$

For $R_{5,1}$, by using a similar argument as in the proof for R_1, we have

$$R_{5,1} = \frac{1}{\Sigma} \sum_{g\in[G]} \mathbb{E}\left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]}\right) \left(\bar{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right) + o_P(1).$$

In addition, by using a similar argument as in the proof for $R_{3,4}$, we have $R_{5,2} = o_P(1)$ and

$R_{5,3} = o_P(1)$. Lastly, by using a similar argument as in the proof for $R_{2,2}$, we have

$$R_{5,4} = \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right) + o_P(1).$$

Combining the results with (C.13), we have

$$\begin{aligned}
&\frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} X_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{X}_{[k]}^\top Q_{[k,g]} e_{[g]} \right) \\
&= \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]} \right) \left(\bar{\Pi}_{[g]}^\top \tilde{e}_{[g]} \right) \\
&+ \frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right) + o_P(1).
\end{aligned}$$

This concludes the proof. □

C.6 Proof of Lemma B.6

For the first result in Lemma B.6, we note that

$$\begin{aligned}
&\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) \\
&= \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)}_{R_6} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)}_{R_7} \\
&+ 2 \times \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)}_{R_8}.
\end{aligned}$$

For R_6, it has mean zero and

$$\begin{aligned}
\mathbb{V}(R_6) &= \frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\tilde{\Pi}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(Q_{[g,h]}Q_{[h,g]}\right) \\
&= o(1),
\end{aligned}$$

whence $R_6 = o_P(1)$. For R_7, since $K/\Sigma = O(1)$, by an application of Lemma B.4, we have

$$R_7 = \frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right) + o_P(1).$$

For R_8, we have

$$\begin{aligned}
\mathbb{V}(R_8) &= \frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^\top - \Omega_h^{\tilde{e},\tilde{V}}\right)Q_{[h,g]}\tilde{e}_{[g]}\right)^2 \\
&+ \frac{C}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\Omega_h^{\tilde{e},\tilde{V}}Q_{[h,g]}\tilde{e}_{[g]}\right)^2 \\
&= o(1),
\end{aligned}$$

where the last equality holds because

$$\begin{aligned}
&\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^\top - \Omega_h^{\tilde{e},\tilde{V}}\right)Q_{[h,g]}\tilde{e}_{[g]}\right)^2 \\
\leqslant &\frac{C}{\Sigma^2}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\left(\tilde{e}_{[h]}\tilde{V}_{[h]}^\top - \Omega_h^{\tilde{e},\tilde{V}}\right)Q_{[h,g]}\tilde{e}_{[g]}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \frac{C}{\Sigma^2} \sum_{g,h \in [G]^2, g \neq h} \operatorname{trace}\left(Q_{[g,h]} Q_{[h,g]}\right) \\
&= o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{\Sigma^2} \mathbb{E}\left(\sum_{g,h \in [G]^2, g \neq h} \tilde{\Pi}_{[g]}^{\top} Q_{[g,h]} \Omega_h^{\tilde{e},\tilde{V}} Q_{[h,g]} \tilde{e}_{[g]}\right)^2 \\
&= \frac{1}{\Sigma^2} \mathbb{E}\left(\sum_{g \in [G]} \tilde{\Pi}_{[g]}^{\top} \left(\sum_{h \neq g} Q_{[g,h]} \Omega_h^{\tilde{e},\tilde{V}} Q_{[h,g]}\right) \tilde{e}_{[g]}\right)^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g \in [G]} \operatorname{trace}\left(\left(\sum_{h \neq g} Q_{[g,h]} \Omega_h^{\tilde{e},\tilde{V}} Q_{[h,g]}\right)\left(\sum_{k \neq g} Q_{[g,k]} \Omega_k^{\tilde{V},\tilde{e}} Q_{[k,g]}\right)\right) \\
&\leqslant \frac{C}{\Sigma^2} \operatorname{trace}\left((Q - \bar{Q}) \Omega^{\tilde{e},\tilde{V}} (Q - \bar{Q})^2 \Omega^{\tilde{V},\tilde{e}} (Q - \bar{Q})\right) \\
&= o(1).
\end{aligned}
$$

It follows that $R_8 = o_P(1)$. Combining the results above, we have

$$
\begin{aligned}
&\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^{\top} Q_{[g,h]} \tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top} Q_{[h,g]} \tilde{e}_{[g]}\right) \\
&= \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left(\tilde{V}_{[g]}^{\top} Q_{[g,h]} \tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^{\top} Q_{[h,g]} \tilde{e}_{[g]}\right) + o_P(1).
\end{aligned}
$$

Next, we turn to the second result in Lemma B.6. We note that

$$
\begin{aligned}
&\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^{\top} Q_{[g,h]} e_{[h]}\right)\left(\tilde{X}_{[h]}^{\top} Q_{[h,g]} e_{[g]}\right) \\
&= \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^{\top} Q_{[g,h]} \tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top} Q_{[h,g]} \tilde{e}_{[g]}\right) \\
&+ \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^{\top} Q_{[g,h]} W_{[h]} \hat{\gamma}_{\tilde{e}}\right)\left(\tilde{X}_{[h]}^{\top} Q_{[h,g]} W_{[g]} \hat{\gamma}_{\tilde{e}}\right) \\
&- \frac{2}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^{\top} Q_{[g,h]} W_{[h]} \hat{\gamma}_{\tilde{e}}\right)\left(\tilde{X}_{[h]}^{\top} Q_{[h,g]} \tilde{e}_{[g]}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
&\left|\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}W_{[h]}\hat{\gamma}_{\tilde{e}}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)\right| \\
\leqslant\ &\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}W_{[h]}\hat{\gamma}_{\tilde{e}}\right)^2 \\
\leqslant\ &\max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2^2\times\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left\|Q_{[h,g]}\tilde{X}_{[g]}\right\|_2^2 \\
=\ &o_P(1),
\end{aligned}
$$

by Lemma B.1, and

$$
\begin{aligned}
&\left|\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}W_{[h]}\hat{\gamma}_{\tilde{e}}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\right| \\
\leqslant\ &\left(\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}W_{[h]}\hat{\gamma}_{\tilde{e}}\right)^2\right)^{1/2} \\
\times\ &\left(\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2\right)^{1/2} \\
=\ &o_P(1).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}e_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}e_{[g]}\right) \\
=\ &\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)+o_P(1) \\
=\ &\frac{1}{\Sigma}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)+o_P(1).
\end{aligned}
$$

Next, we turn to the third result in Lemma B.6. Note that

$$
\begin{aligned}
&\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} X_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} X_{[g]} \right) \\
&= \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} V_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} V_{[g]} \right)}_{R_9} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \Pi_{[g]} \right)}_{R_{10}} \\
&+ 2 \times \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} V_{[g]} \right)}_{R_{11}} .
\end{aligned}
$$

By using the same argument as in the proof above, we have

$$
R_9 = \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) + o_P(1).
$$

For R_{10}, we have

$$
\begin{aligned}
R_{10} &= \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \Pi_{[g]} \right) \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \Pi_{[g]} \right)}_{R_{10,1}} \\
&+ 2 \times \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \Pi_{[g]} \right)}_{R_{10,2}} .
\end{aligned}
$$

Note that

$$
\mathbb{V}\left(R_{10,1}\right) = \frac{1}{\Sigma^2} \mathbb{E} \left(\sum_{g,h \in [G]^2, g \neq h} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \Pi_{[g]} \right) \right)^2
$$

$$
\begin{aligned}
&\leqslant \frac{C}{\Sigma^2} \sum_{g,h\in[G]^2, g\neq h} \mathbb{E}\left(\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)\right)^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g,h\in[G]^2, g\neq h} \operatorname{trace}\left(Q_{[h,g]}Q_{[g,h]}\right) \\
&= o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{V}\left(R_{10,2}\right) &= \frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g,h\in[G]^2, g\neq h}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right)\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\left\|\sum_{g\in[G], g\neq h} Q_{[h,g]}\Pi_{[g]}\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right\|_2^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\left\|\sum_{g\in[G], g\neq h} Q_{[h,g]}\Pi_{[g]}\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\right\|_F^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\sum_{g,g'\in[G]^2, g,g'\neq h} \operatorname{trace}\left(Q_{[h,g]}\Pi_{[g]}\tilde{\Pi}_{[g]}^\top Q_{[g,h]}Q_{[h,g']}\tilde{\Pi}_{[g']}\Pi_{[g']}^\top Q_{[g',h]}\right) \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\sum_{g,g'\in[G]^2, g,g'\neq h}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}Q_{[h,g']}\tilde{\Pi}_{[g']}\right)^2 \\
&+ \frac{C}{\Sigma^2}\sum_{h\in[G]}\sum_{g,g'\in[G]^2, g,g'\neq h}\left(\Pi_{[g]}^\top Q_{[g,h]}Q_{[h,g']}\Pi_{[g']}\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h\in[G]}\sum_{g,g'\in[G]^2, g,g'\neq h} \operatorname{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)\operatorname{trace}\left(Q_{[h,g']}Q_{[g',h]}\right) \\
&= o(1).
\end{aligned}
$$

Therefore, we have

$$
R_{10} = \frac{1}{\Sigma}\sum_{g,h\in[G]^2, g\neq h}\left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]}\Pi_{[g]}\right) + o_P(1).
$$

For R_{11}, we have

$$
\begin{aligned}
R_{11} &= \frac{1}{\Sigma} \sum_{g,h\in[G]^2, g\neq h} \left(\tilde{X}_{[g]}^{\top} Q_{[g,h]} \Pi_{[h]}\right) \left(\tilde{X}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]}\right) + o_P(1) \\
&= \underbrace{\frac{1}{\Sigma} \sum_{g,h\in[G]^2, g\neq h} \left(\tilde{\Pi}_{[g]}^{\top} Q_{[g,h]} \Pi_{[h]}\right) \left(\tilde{\Pi}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]}\right)}_{R_{11,1}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h\in[G]^2, g\neq h} \left(\tilde{\Pi}_{[g]}^{\top} Q_{[g,h]} \Pi_{[h]}\right) \left(\tilde{V}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]}\right)}_{R_{11,2}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h\in[G]^2, g\neq h} \left(\tilde{V}_{[g]}^{\top} Q_{[g,h]} \Pi_{[h]}\right) \left(\tilde{\Pi}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]}\right)}_{R_{11,3}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h\in[G]^2, g\neq h} \left(\tilde{V}_{[g]}^{\top} Q_{[g,h]} \Pi_{[h]}\right) \left(\tilde{V}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]}\right)}_{R_{11,4}}.
\end{aligned}
$$

By using a similar argument as in the proof for $R_{10,1}$ and $R_{10,2}$, we have $R_{11,1} = o_P(1)$ and $R_{11,2} = o_P(1)$. For $R_{11,3}$, we have

$$
\begin{aligned}
\mathbb{V}\left(R_{11,3}\right) &= \frac{1}{\Sigma^2} \mathbb{E}\left(\sum_{g,h\in[G]^2, g\neq h} \left(\tilde{V}_{[g]}^{\top} Q_{[g,h]} \Pi_{[h]}\right) \left(\tilde{\Pi}_{[h]}^{\top} Q_{[h,g]} \tilde{V}_{[g]}\right)\right)^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g\in[G]} \left\| \sum_{h\in[G], h\neq g} Q_{[g,h]} \Pi_{[h]} \tilde{\Pi}_{[h]}^{\top} Q_{[h,g]} \right\|_F^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g\in[G]} \sum_{h,h'\in[G]^2, h,h'\neq g} \left(\tilde{\Pi}_{[h]}^{\top} Q_{[h,g]} Q_{[g,h']} \tilde{\Pi}_{[h']}\right)^2 \\
&+ \frac{C}{\Sigma^2} \sum_{g\in[G]} \sum_{h,h'\in[G]^2, h,h'\neq g} \left(\Pi_{[h]}^{\top} Q_{[h,g]} Q_{[g,h']} \Pi_{[h']}\right)^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g\in[G]} \sum_{h,h'\in[G]^2, h,h'\neq g} \operatorname{trace}\left(Q_{[g,h]} Q_{[h,g]}\right) \operatorname{trace}\left(Q_{[g,h']} Q_{[h',g]}\right) \\
&= o(1),
\end{aligned}
$$

and thus

$$R_{11,3} = \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) + o_P(1).$$

For $R_{11,4}$, we have

$$\begin{aligned}
\mathbb{V}\left(R_{11,4}\right) \leqslant \mathbb{V}&\left(\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[h]}^\top Q_{[h,g]} \Omega_g^{\tilde{V},\tilde{V}} Q_{[g,h]} \Pi_{[h]} \right) \\
&+ \mathbb{V}\left(\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[h]}^\top Q_{[h,g]} (\tilde{V}_{[g]} \tilde{V}_{[g]}^\top - \Omega_g^{\tilde{V},\tilde{V}}) Q_{[g,h]} \Pi_{[h]} \right),
\end{aligned}$$

where

$$\begin{aligned}
&\mathbb{V}\left(\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[h]}^\top Q_{[h,g]} \Omega_g^{\tilde{V},\tilde{V}} Q_{[g,h]} \Pi_{[h]} \right) \\
&= \frac{1}{\Sigma^2} \mathbb{E}\left(\sum_{h \in [G]} \tilde{V}_{[h]}^\top \left(\sum_{g \in [G], g \neq h} Q_{[h,g]} \Omega_g^{\tilde{V},\tilde{V}} Q_{[g,h]} \right) \Pi_{[h]} \right)^2 \\
&\leqslant \frac{C \Pi^\top \Pi}{(\Pi^\top \Pi + K)^2} \\
&= o(1),
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{V}\left(\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[h]}^\top Q_{[h,g]} (\tilde{V}_{[g]} \tilde{V}_{[g]}^\top - \Omega_g^{\tilde{V},\tilde{V}}) Q_{[g,h]} \Pi_{[h]} \right) \\
&= \frac{1}{\Sigma^2} \mathbb{E}\left(\sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[h]}^\top Q_{[h,g]} (\tilde{V}_{[g]} \tilde{V}_{[g]}^\top - \Omega_g^{\tilde{V},\tilde{V}}) Q_{[g,h]} \Pi_{[h]} \right)^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right)^2 \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g,h \in [G]^2, g \neq h} \operatorname{trace}\left(Q_{[h,g]} Q_{[g,h]} \right)
\end{aligned}$$

$$= o(1).$$

It follows that $R_{11,4} = o_P(1)$, whence

$$R_{11} = \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) + o_P(1).$$

Combining the results above, we have

$$\begin{aligned}
&\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} X_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} X_{[g]} \right) \\
&= \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \\
&+ \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{\Pi}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \Pi_{[g]} \right) \\
&+ \frac{2}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) + o_P(1).
\end{aligned}$$

Finally, for the last result of Lemma B.6, we note that

$$\begin{aligned}
&\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} X_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} e_{[g]} \right) \\
&= \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} V_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} e_{[g]} \right)}_{R_{12}} \\
&+ \underbrace{\frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} e_{[g]} \right)}_{R_{13}}.
\end{aligned}$$

By using the same argument as in the proof for R_9, we have

$$R_{12} = \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) + o_P(1).$$

In addition, by using the same argument as in the proof for R_{11}, we have

$$R_{13} = \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) + o_P(1).$$

It follows that

$$\begin{aligned} & \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} X_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} e_{[g]} \right) \\ &= \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) \\ &+ \frac{1}{\Sigma} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]} \Pi_{[h]} \right) \left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) + o_P(1). \end{aligned}$$

This concludes the proof. □

C.7 Proof of Lemma B.7

Throughout the proof we denote $\acute{\Delta} = \acute{\beta} - \beta$, and note that $\acute{\Delta} = o_P(1)$ by assumption. We divide the proof into four steps.

Step 1: Consistency of $\acute{\Omega}$. Note that $1/C \leqslant \lambda_{\min}(\Omega/n) \leqslant \lambda_{\min}(\Omega/n) \leqslant C$ by Assumption 1, and thus it suffices to show that

$$\frac{1}{n}\acute{\Omega} - \frac{1}{n}\Omega = o_P(1).$$

Let

$$\tilde{\Omega} = \sum_{g \in [G]} \left(z_{[g]}^\top \tilde{e}_{[g]} \right) \left(z_{[g]}^\top \tilde{e}_{[g]} \right)^\top \quad \text{and} \quad \bar{\Omega} = \sum_{g \in [G]} \left(z_{[g]}^\top e_{[g]} \right) \left(z_{[g]}^\top e_{[g]} \right)^\top.$$

We aim to show that

$$\frac{1}{n}\tilde{\Omega} - \frac{1}{n}\Omega = o_P(1), \tag{C.14}$$
$$\frac{1}{n}\bar{\Omega} - \frac{1}{n}\tilde{\Omega} = o_P(1), \tag{C.15}$$
$$\frac{1}{n}\acute{\Omega} - \frac{1}{n}\bar{\Omega} = o_P(1). \tag{C.16}$$

For (C.14), consider its (j,k)-th element for $1 \leqslant j,k \leqslant d_z$, given by

$$\frac{1}{n}\sum_{g\in[G]}\left[z_{[g],j}^{\top}\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}z_{[g],k} - \mathbb{E}\left(z_{[g],j}^{\top}\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}z_{[g],k}\right)\right],$$

where we use $z_{[g],j}$ ($z_{[g],k}$) to denote the j-th (k-th) column of $z_{[g]}$; note that it has mean zero and

$$\begin{aligned}
&\mathbb{V}\left(\frac{1}{n}\sum_{g\in[G]}\left[z_{[g],j}^{\top}\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}z_{[g],k} - \mathbb{E}\left(z_{[g],j}^{\top}\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}z_{[g],k}\right)\right]\right)\\
&\leqslant \frac{1}{n^2}\sum_{g\in[G]}\mathbb{E}\left(z_{[g],j}^{\top}\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}z_{[g],k}\right)^2\\
&\leqslant \frac{1}{n^2}\sum_{g\in[G]}z_{[g],j}^{\top}z_{[g],j}z_{[g],k}^{\top}z_{[g],k}\mathbb{E}\left(\tilde{e}_{[g]}^{\top}\tilde{e}_{[g]}\right)^2\\
&\leqslant \frac{C\max_{i\in I_g,g\in[G]}\|z_{i,g}\|_2^2}{n}\times\frac{1}{n}\sum_{g\in[G]}z_{[g],j}^{\top}z_{[g],j}\\
&= o(1)
\end{aligned}$$

by Assumption 1, and the result follows since d_z is fixed.

For (C.15), its (j,k)-th element can be written as

$$\begin{aligned}
&\frac{1}{n}\sum_{g\in[G]}\left(z_{[g],j}^{\top}e_{[g]}e_{[g]}^{\top}z_{[g],k} - z_{[g],j}^{\top}\tilde{e}_{[g]}\tilde{e}_{[g]}^{\top}z_{[g],k}\right)\\
&= \frac{1}{n}\sum_{g\in[G]}(z_{[g],j}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}})(z_{[g],k}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}})
\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{n} \sum_{g \in [G]} (z_{[g],j}^\top W_{[g]} \hat{\gamma}_{\tilde{e}})(z_{[g],k}^\top \tilde{e}_{[g]}) \\
&- \frac{1}{n} \sum_{g \in [G]} (z_{[g],k}^\top W_{[g]} \hat{\gamma}_{\tilde{e}})(z_{[g],j}^\top \tilde{e}_{[g]}).
\end{aligned}
$$

We have

$$
\begin{aligned}
&\left| \frac{1}{n} \sum_{g \in [G]} (z_{[g],j}^\top W_{[g]} \hat{\gamma}_{\tilde{e}})(z_{[g],k}^\top W_{[g]} \hat{\gamma}_{\tilde{e}}) \right| \\
&\leqslant \sqrt{\frac{1}{n} \sum_{g \in [G]} (z_{[g],j}^\top W_{[g]} \hat{\gamma}_{\tilde{e}})^2} \times \sqrt{\frac{1}{n} \sum_{g \in [G]} (z_{[g],k}^\top W_{[g]} \hat{\gamma}_{\tilde{e}})^2} \\
&\leqslant \max_{1 \leqslant g \leqslant G} \left\| W_{[g]} \hat{\gamma}_{\tilde{e}} \right\|_2^2 \times \sqrt{\frac{1}{n} \sum_{g \in [G]} z_{[g],j}^\top z_{[g],j}} \times \sqrt{\frac{1}{n} \sum_{g \in [G]} z_{[g],k}^\top z_{[g],k}} \\
&\leqslant \max_{1 \leqslant g \leqslant G} \left\| W_{[g]} \hat{\gamma}_{\tilde{e}} \right\|_2^2 \times \sqrt{v_j^\top \left(\frac{1}{n} \sum_{i \in I_g, g \in [G]} z_{i,g} z_{i,g}^\top \right) v_j} \times \sqrt{v_k^\top \left(\frac{1}{n} \sum_{i \in I_g, g \in [G]} z_{i,g} z_{i,g}^\top \right) v_k} \\
&= o_P(1)
\end{aligned}
$$

by Assumption 1 and Lemma B.1, where we use v_j (v_k) to denote the d_z-dimensional unit vector with j-th (k-th) element one and other elements zero; we also have

$$
\begin{aligned}
&\left| \frac{1}{n} \sum_{g \in [G]} (z_{[g],j}^\top W_{[g]} \hat{\gamma}_{\tilde{e}})(z_{[g],k}^\top \tilde{e}_{[g]}) \right| \\
&\leqslant \max_{1 \leqslant g \leqslant G} \left\| W_{[g]} \hat{\gamma}_{\tilde{e}} \right\|_2 \times \sqrt{\frac{1}{n} \sum_{g \in [G]} z_{[g],j}^\top z_{[g],j}} \times \sqrt{\frac{1}{n} \sum_{g \in [G]} z_{[g],k}^\top \tilde{e}_{[g]} \tilde{e}_{[g]}^\top z_{[g],k}} \\
&= o_P(1),
\end{aligned}
$$

and by using the same argument,

$$
\left| \frac{1}{n} \sum_{g \in [G]} (z_{[g],k}^\top W_{[g]} \hat{\gamma}_{\tilde{e}})(z_{[g],j}^\top \tilde{e}_{[g]}) \right| = o_P(1).
$$

For (C.16), its (j,k)-th element can be written as

$$\begin{aligned}
&\frac{1}{n}\sum_{g\in[G]}\left(z_{[g],j}^{\top}\acute{e}_{[g]}\acute{e}_{[g]}^{\top}z_{[g],k}-z_{[g],j}^{\top}e_{[g]}e_{[g]}^{\top}z_{[g],k}\right)\\
&=\frac{\acute{\Delta}^{2}}{n}\sum_{g\in[G]}(z_{[g],j}^{\top}X_{[g]})(z_{[g],k}^{\top}X_{[g]})\\
&-\frac{\acute{\Delta}}{n}\sum_{g\in[G]}(z_{[g],j}^{\top}X_{[g]})(z_{[g],k}^{\top}e_{[g]})\\
&-\frac{\acute{\Delta}}{n}\sum_{g\in[G]}(z_{[g],k}^{\top}X_{[g]})(z_{[g],j}^{\top}e_{[g]}),
\end{aligned}$$

and note that

$$\begin{aligned}
\frac{1}{n}\sum_{g\in[G]}(z_{[g],j}^{\top}X_{[g]})(z_{[g],k}^{\top}X_{[g]}) &\leqslant \sqrt{\frac{1}{n}\sum_{g\in[G]}z_{[g],j}^{\top}z_{[g],j}X_{[g]}^{\top}X_{[g]}}\times\sqrt{\frac{1}{n}\sum_{g\in[G]}z_{[g],k}^{\top}z_{[g],k}X_{[g]}^{\top}X_{[g]}}\\
&=O_P(1),
\end{aligned}$$

since $\max_{g\in[G]}\mathbb{E}\left(X_{[g]}^{\top}X_{[g]}\right)=O(1)$ by Lemma B.1; the other two terms can be handled similarly.

Step 2: Consistency of $\acute{\Psi}$. By Assumption 2, we have

$$\frac{1}{r_n}z^{\top}X=\frac{1}{r_n}z^{\top}\Pi+\frac{1}{r_n}z^{\top}V=\frac{1}{r_n}z^{\top}\Pi+o_P(1), \tag{C.17}$$

and note that

$$\begin{aligned}
\frac{1}{\lambda_n}\hat{A}_n &= \frac{1}{\lambda_n}A_n+\frac{1}{\lambda_n}(\hat{A}_n-A_n)\\
&=\frac{1}{\lambda_n}A_n+(\frac{1}{\lambda_n}A_n)^{1/2}(A_n^{-1/2}\hat{A}_nA_n^{-1/2}-I_{d_z})(\frac{1}{\lambda_n}A_n)^{1/2}\\
&=\frac{1}{\lambda_n}A_n+o_P(1),
\end{aligned} \tag{C.18}$$

where $\lambda_n = \lambda_{\max}(A_n)$. In addition, we have

$$\frac{1}{(\frac{1}{r_n}\Pi^\top z)(\frac{1}{\lambda_n}A_n)(\frac{1}{n}\Omega)(\frac{1}{\lambda_n}A_n)(\frac{1}{r_n}z^\top\Pi)} \leqslant C.$$

Therefore, we have

$$\frac{\acute{\Psi}}{\Psi} = \frac{X^\top z\hat{A}_n\acute{\Omega}\hat{A}_n z^\top X}{\Pi^\top z A_n \Omega A_n z^\top \Pi} = \frac{(\frac{1}{r_n}X^\top z)(\frac{1}{\lambda_n}\hat{A}_n)(\frac{1}{n}\acute{\Omega})(\frac{1}{\lambda_n}\hat{A}_n)(\frac{1}{r_n}z^\top X)}{(\frac{1}{r_n}\Pi^\top z)(\frac{1}{\lambda_n}A_n)(\frac{1}{n}\Omega)(\frac{1}{\lambda_n}A_n)(\frac{1}{r_n}z^\top\Pi)} = 1 + o_P(1)$$

by Step 1.

Step 3: Consistency of $\acute{\Sigma}$. Define

$$\bar{\Sigma} = \mathbb{E}\left(\tilde{\Pi}^\top(Q-\bar{Q})\tilde{e}\right)^2 + \mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right)^2,$$

and recall

$$\Sigma = \mathbb{E}\left(\hat{\Pi}^\top\tilde{e}\right)^2 + \mathbb{E}\left(\tilde{V}^\top(P-\bar{P})\tilde{e}\right)^2 = \mathbb{E}\left(\tilde{\Pi}^\top Q\tilde{e}\right)^2 + \mathbb{E}\left(\tilde{V}^\top(P-\bar{P})\tilde{e}\right)^2.$$

Using the notation $\|X\|_{\mathbb{P},2} = \left(\mathbb{E}\left(X^2\right)\right)^{1/2}$, we have

$$\begin{aligned}
\left|\frac{\bar{\Sigma}-\Sigma}{\Sigma}\right| &\leqslant \left|\frac{\left\|\tilde{\Pi}^\top Q\tilde{e}\right\|_{\mathbb{P},2}^2 - \left\|\tilde{\Pi}^\top(Q-\bar{Q})\tilde{e}\right\|_{\mathbb{P},2}^2}{\Sigma}\right| + \left|\frac{\left\|\tilde{V}^\top(P-\bar{P})\tilde{e}\right\|_{\mathbb{P},2}^2 - \left\|\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right\|_{\mathbb{P},2}^2}{\Sigma}\right| \\
&\leqslant \left(\frac{2\left\|\tilde{\Pi}^\top Q\tilde{e}\right\|_{\mathbb{P},2}\left\|\tilde{\Pi}^\top\bar{Q}\tilde{e}\right\|_{\mathbb{P},2} + \left\|\tilde{\Pi}^\top\bar{Q}\tilde{e}\right\|_{\mathbb{P},2}^2}{\Sigma}\right) + \left|\frac{\left\|\tilde{V}^\top(P-\bar{P})\tilde{e}\right\|_{\mathbb{P},2}^2 - \left\|\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right\|_{\mathbb{P},2}^2}{\Sigma}\right| \\
&\leqslant \left(\frac{2\left\|\tilde{\Pi}^\top Q\tilde{e}\right\|_{\mathbb{P},2}\left\|\tilde{\Pi}^\top\bar{Q}\tilde{e}\right\|_{\mathbb{P},2} + \left\|\tilde{\Pi}^\top\bar{Q}\tilde{e}\right\|_{\mathbb{P},2}^2}{\Pi^\top\Pi}\right) + \left|\frac{\left\|\tilde{V}^\top(P-\bar{P})\tilde{e}\right\|_{\mathbb{P},2}^2 - \left\|\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right\|_{\mathbb{P},2}^2}{K}\right| \\
&= o(1),
\end{aligned}$$

where the first inequality is by triangular inequality, the second inequality is by Cauchy-Schwarz inequality, the third inequality is by Lemma B.5, and the last equality holds because

$$\mathbb{E}\left(\tilde{\Pi}^\top \bar{Q}\tilde{e}\right)^2 \leqslant C\left\|\bar{Q}\right\|_{op}^2 \tilde{\Pi}^\top\tilde{\Pi} = o(\Pi^\top\Pi)$$

by Lemma B.2, and

$$\begin{aligned}\frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right)^2 &= \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left[\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)^2 + \left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\right]\\ &= \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left[\left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)^2 + \left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)\right] + o(1)\\ &= \frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(P-\bar{P})\tilde{e}\right)^2 + o(1)\end{aligned}$$

by Lemma B.4. Therefore, we have

$$\frac{\bar{\Sigma}}{\Sigma} \to 1.$$

In addition, note that we can write

$$\bar{\Sigma} = \sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2 + \sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right),$$

as in Chao et al. (2012), and thus

$$\begin{aligned}\frac{\acute{\Sigma}-\bar{\Sigma}}{\bar{\Sigma}} &= \underbrace{\frac{1}{\bar{\Sigma}}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2 - \sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2\right)}_{R_{14}}\\ &\quad+\underbrace{\frac{1}{\bar{\Sigma}}\left(\begin{array}{c}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\\ -\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\end{array}\right)}_{R_{15}}\end{aligned}$$

$$+\underbrace{\frac{1}{\bar{\bar{\Sigma}}}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\acute{e}_{[g]}\right)^2-\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2\right)}_{R_{16}}$$

$$+\underbrace{\frac{1}{\bar{\bar{\Sigma}}}\left(\begin{array}{c}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}\acute{e}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}\acute{e}_{[g]}\right)\\ -\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)\end{array}\right)}_{R_{17}}.$$

Note also that, by Lemmas B.5 and B.6, we have $R_{14}=o_P(1)$ and $R_{15}=o_P(1)$.

For R_{16}, we have

$$R_{16}=\underbrace{\frac{1}{\bar{\bar{\Sigma}}}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\acute{e}_{[g]}\right)^2-\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right)^2\right)}_{R_{16,1}}$$

$$+\underbrace{\frac{1}{\bar{\bar{\Sigma}}}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right)^2-\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right)^2\right)}_{R_{16,2}}.$$

For $R_{16,1}$, we have

$$\begin{aligned}R_{16,1}=&\frac{\acute{\Delta}^2}{\bar{\bar{\Sigma}}}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right)^2\\&-\frac{2\acute{\Delta}}{\bar{\bar{\Sigma}}}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^{\top}Q_{[k,g]}e_{[g]}\right).\end{aligned}$$

By Lemma B.5, we have

$$\frac{1}{\bar{\bar{\Sigma}}}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right)^2=O_P(1),$$

since

$$\frac{1}{\bar{\bar{\Sigma}}} \sum_{g \in [G]} \left(\bar{\Pi}_{[g]}^\top \Pi_{[g]} \right)^2 \leqslant \frac{C \bar{\Pi}^\top \bar{\Pi}}{\bar{\Sigma}} = O(1),$$

$$\frac{1}{\bar{\bar{\Sigma}}} \sum_{g \in [G]} \mathbb{E} \left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]} \right)^2 \leqslant \frac{C \bar{\Pi}^\top \bar{\Pi}}{\bar{\Sigma}} = O(1),$$

$$\frac{1}{\bar{\bar{\Sigma}}} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \Pi_{[g]} \right)^2 \leqslant \frac{C \Pi^\top \Pi}{\bar{\Sigma}} = O(1),$$

$$\frac{1}{\bar{\bar{\Sigma}}} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right)^2 \leqslant \frac{C}{\bar{\bar{\Sigma}}} \sum_{g,h \in [G]^2, g \neq h} \operatorname{trace} \left(Q_{[g,h]} Q_{[h,g]} \right) = O(1).$$

By the same lemma, we also have

$$\frac{1}{\bar{\bar{\Sigma}}} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} X_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{X}_{[k]}^\top Q_{[k,g]} e_{[g]} \right) = O_P(1),$$

since

$$\left| \frac{1}{\bar{\bar{\Sigma}}} \sum_{g \in [G]} \mathbb{E} \left(\bar{\Pi}_{[g]}^\top \tilde{V}_{[g]} \right) \left(\bar{\Pi}_{[g]}^\top \tilde{e}_{[g]} \right) \right| \leqslant \frac{C \bar{\Pi}^\top \bar{\Pi}}{\bar{\Sigma}} = O(1)$$

and

$$\left| \frac{1}{\bar{\bar{\Sigma}}} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right) \right|$$
$$\leqslant \frac{C}{\bar{\bar{\Sigma}}} \sum_{g,h \in [G]^2, g \neq h} \operatorname{trace} \left(Q_{[g,h]} Q_{[h,g]} \right) = O(1).$$

It follows that $R_{16,1} = o_P(1)$, and since $R_{16,2} = o_P(1)$ by Lemma B.5, we have $R_{16} = o_P(1)$.

For R_{17}, we have

$$R_{17} = \underbrace{\frac{1}{\bar{\bar{\Sigma}}}\left(\begin{array}{c} \sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}\acute{e}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}\acute{e}_{[g]}\right) \\ -\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}e_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right) \end{array}\right)}_{R_{17,1}} + \underbrace{\frac{1}{\bar{\bar{\Sigma}}}\left(\begin{array}{c} \sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}e_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right) \\ -\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{e}_{[g]}\right) \end{array}\right)}_{R_{17,2}}.$$

For $R_{17,1}$, we have

$$\begin{aligned} R_{17,1} &= \frac{\hat{\Delta}^2}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}X_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right) \\ &\quad - \frac{2\hat{\Delta}}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}X_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}e_{[g]}\right). \end{aligned}$$

By Lemma B.6, we have

$$\frac{1}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}X_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}X_{[g]}\right) = O_P(1),$$

since

$$\begin{aligned} &\left|\frac{1}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{X}_{[g]}^{\top}Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{X}_{[h]}^{\top}Q_{[h,g]}\tilde{V}_{[g]}\right)\right| \leqslant \frac{C}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left(Q_{[h,g]}Q_{[g,h]}\right) = O(1), \\ &\left|\frac{1}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{\Pi}_{[g]}^{\top}Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\Pi_{[g]}\right)\right| \leqslant \frac{C\Pi^{\top}\Pi}{\bar{\bar{\Sigma}}} = O(1), \\ &\left|\frac{1}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^{\top}Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{\Pi}_{[h]}^{\top}Q_{[h,g]}\tilde{V}_{[g]}\right)\right| \leqslant \frac{C\Pi^{\top}\Pi}{\bar{\bar{\Sigma}}} = O(1). \end{aligned}$$

By the same lemma, we also have

$$\frac{1}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}X_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}e_{[g]}\right)=O_P(1),$$

since

$$\left|\frac{1}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\right|\leqslant\frac{C}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\text{trace}\left(Q_{[h,g]}Q_{[g,h]}\right)=O(1),$$
$$\left|\frac{1}{\bar{\bar{\Sigma}}}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\Pi_{[h]}\right)\left(\tilde{\Pi}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\right|\leqslant\frac{C\Pi^\top\Pi}{\bar{\Sigma}}=O(1).$$

It follows that $R_{17,1}=o_P(1)$, and since $R_{17,2}=o_P(1)$ by Lemma B.6, we have $R_{17}=o_P(1)$. Combining the results above, we have

$$\frac{\acute{\Sigma}-\bar{\Sigma}}{\bar{\Sigma}}=o_P(1),$$

and the desired result follows from $\Sigma/\bar{\Sigma}\to 1$.

Step 4: Consistency of $\acute{\Upsilon}$. We have

$$\begin{aligned}\frac{\acute{\Upsilon}-\Upsilon}{\Upsilon}&=2\times\underbrace{\frac{1}{\Upsilon}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)^2-\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)^2\right)}_{R_{18}}\\&+2\times\underbrace{\frac{1}{\Upsilon}\left(\sum_{g,h\in[G]^2,g\neq h}\left(e_{[g]}^\top P_{[g,h]}e_{[h]}\right)^2-\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)^2\right)}_{R_{19}}\\&+2\times\underbrace{\frac{1}{\Upsilon}\left(\sum_{g,h\in[G]^2,g\neq h}\left(\acute{e}_{[g]}^\top P_{[g,h]}\acute{e}_{[h]}\right)^2-\sum_{g,h\in[G]^2,g\neq h}\left(e_{[g]}^\top P_{[g,h]}e_{[h]}\right)^2\right)}_{R_{20}},\end{aligned}$$

and note that

$$\begin{aligned}
\Upsilon &= \mathbb{V}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{e}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\\
&= \mathbb{E}\left(\sum_{g,h\in[G]^2,g\neq h}\tilde{e}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)^2\\
&= 2\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{e}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)^2\\
&= 2\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left[\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\Omega_h^{\tilde{e},\tilde{e}}P_{[h,g]}\right]\\
&\geqslant \frac{1}{C}\sum_{g,h\in[G]^2,g\neq h}\operatorname{trace}\left[P_{[g,h]}P_{[h,g]}\right]\\
&\geqslant \frac{1}{C}K. \qquad\qquad (\text{C.19})
\end{aligned}$$

By Lemma B.4, we have $R_{18}=o_P(1)$ and $R_{19}=o_P(1)$. For R_{20}, it can be written as

$$\begin{aligned}
R_{20} &= \frac{1}{\Upsilon}\sum_{g,h\in[G]^2,g\neq h}\Big(\big((\acute{e}_{[g]}-e_{[g]})^{\top}P_{[g,h]}(\acute{e}_{[h]}-e_{[h]})+e_{[g]}^{\top}P_{[g,h]}(\acute{e}_{[h]}-e_{[h]})\\
&\qquad +(\acute{e}_{[g]}-e_{[g]})^{\top}P_{[g,h]}e_{[h]}+e_{[g]}^{\top}P_{[g,h]}e_{[h]}\big)^2-\left(e_{[g]}^{\top}P_{[g,h]}e_{[h]}\right)^2\Big)\\
&= \acute{\Delta}^4\times\underbrace{\frac{1}{\Upsilon}\sum_{g,h\in[G]^2,g\neq h}\left(X_{[g]}^{\top}P_{[g,h]}X_{[h]}\right)^2}_{R_{20,1}}\\
&+\acute{\Delta}^2\times\underbrace{\frac{1}{\Upsilon}\sum_{g,h\in[G]^2,g\neq h}\left(e_{[g]}^{\top}P_{[g,h]}X_{[h]}\right)^2}_{R_{20,2}}\\
&+\acute{\Delta}^2\times\underbrace{\frac{1}{\Upsilon}\sum_{g,h\in[G]^2,g\neq h}\left(X_{[g]}^{\top}P_{[g,h]}e_{[h]}\right)^2}_{R_{20,3}}\\
&-4\times\acute{\Delta}^3\times\underbrace{\frac{1}{\Upsilon}\sum_{g,h\in[G]^2,g\neq h}\left(X_{[g]}^{\top}P_{[g,h]}X_{[h]}\right)\left(e_{[g]}^{\top}P_{[g,h]}X_{[h]}\right)}_{R_{20,4}}
\end{aligned}$$

$$
\begin{aligned}
&+ 2 \times \acute{\Delta}^2 \times \underbrace{\frac{1}{\Upsilon} \sum_{g,h \in [G]^2, g \neq h} \left(X_{[g]}^\top P_{[g,h]} X_{[h]} \right) \left(e_{[g]}^\top P_{[g,h]} e_{[h]} \right)}_{R_{20,5}} \\
&+ 2 \times \acute{\Delta}^2 \times \underbrace{\frac{1}{\Upsilon} \sum_{g,h \in [G]^2, g \neq h} \left(e_{[g]}^\top P_{[g,h]} X_{[h]} \right) \left(X_{[g]}^\top P_{[g,h]} e_{[h]} \right)}_{R_{20,6}} \\
&- 4 \times \acute{\Delta} \times \underbrace{\frac{1}{\Upsilon} \sum_{g,h \in [G]^2, g \neq h} \left(e_{[g]}^\top P_{[g,h]} X_{[h]} \right) \left(e_{[g]}^\top P_{[g,h]} e_{[h]} \right)}_{R_{20,7}}.
\end{aligned}
$$

By using the same argument as in the proof of Lemma B.6, we can show that

$$
R_{20,i} = O_P(1), \quad i = 1, \ldots, 7,
$$

whence $R_{20} = o_P(1)$. Combining the results above, we have

$$
\frac{\acute{\Upsilon} - \Upsilon}{\Upsilon} = o_P(1).
$$

This concludes the proof. □

C.8 Proof of Lemma B.8

We first introduce some notation. Denote

$$
\begin{aligned}
\dot{\Phi}_1 &= (X^\top z \hat{A}_n z^\top X)^{-1} (X^\top z \hat{A}_n \dot{\Omega} \hat{A}_n z^\top X) (X^\top z \hat{A}_n z^\top X)^{-1}, \\
\dot{\Omega} &= \sum_{g \in [G]} \left(z_{[g]}^\top \dot{e}_{[g]} \right) \left(z_{[g]}^\top \dot{e}_{[g]} \right)^\top,
\end{aligned}
$$

where $\dot{e} = Y - X\hat{\beta}_1$, and denote $\hat{\Delta}_1 = \hat{\beta}_1 - \beta$. In addition, denote

$$
\ddot{\Phi}_2 = \left(X^\top (P - \bar{P}) X \right)^{-1} \ddot{\Sigma} \left(X^\top (P - \bar{P}) X \right)^{-1}
$$

$$\ddot{\Sigma} = \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} \ddot{e}_{[g]} \right)^2 + \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \ddot{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \ddot{e}_{[g]} \right),$$

where $\ddot{e} = Y - X\hat{\beta}_2$, and denote $\hat{\Delta}_2 = \hat{\beta}_2 - \beta$. Finally, denote $\tilde{\omega}_1 = \dot{\Phi}_2^{1/2} / \left(\dot{\Phi}_1^{1/2} + \ddot{\Phi}_2^{1/2} \right)$ and $\tilde{\omega}_2 = \dot{\Phi}_1^{1/2} / \left(\dot{\Phi}_1^{1/2} + \ddot{\Phi}_2^{1/2} \right)$.

We divide the proof into four steps. In the first step, we show that if Assumptions 1-3 hold, then

$$\hat{\beta}_1 \xrightarrow{p} \beta, \quad \dot{\Phi}_1 = o_P(1), \quad \text{and} \quad (\hat{\beta}_1 - \beta)/\dot{\Phi}_1^{1/2} = O_P(1).$$

In the second step, we show that if Assumptions 1 and 3 hold and $\Pi^\top \Pi / \sqrt{K} \to \infty$, we have

$$(\hat{\beta}_2 - \beta)^2 \Pi^\top \Pi / \sqrt{K} = o_P(1) \quad \text{and} \quad \ddot{\Phi}_2 \Pi^\top \Pi / \sqrt{K} = o_P(1).$$

In the third step, we show that if Assumptions 1-3 hold and $\Pi^\top \Pi / \sqrt{K} = O(1)$, then

$$\hat{\beta}_2 - \beta = O_P(1) \quad \text{and} \quad 1/\ddot{\Phi}_2^{1/2} = O_P(1).$$

In the last step, we show that if Assumptions 1, 2′ and 3 hold, then

$$\hat{\beta}_1 - \beta = O_P(1) \quad \text{and} \quad 1/\dot{\Phi}_1^{1/2} = O_P(1).$$

Consequently, if Assumptions 1-3 hold and $\Pi^\top \Pi / \sqrt{K} \to \infty$, by the results in Steps 1 and 2, we have

$$\begin{aligned}(\hat{\beta} - \beta)^2 \Pi^\top \Pi / \sqrt{K} \leqslant& \left(\frac{\ddot{\Phi}_2^{1/2}}{\dot{\Phi}_1^{1/2} + \ddot{\Phi}_2^{1/2}} \right)^2 \times (\hat{\beta}_1 - \beta)^2 \Pi^\top \Pi / \sqrt{K} \\ &+ \left(\frac{\dot{\Phi}_1^{1/2}}{\dot{\Phi}_1^{1/2} + \ddot{\Phi}_2^{1/2}} \right)^2 \times (\hat{\beta}_2 - \beta)^2 \Pi^\top \Pi / \sqrt{K}\end{aligned}$$

$$\leqslant \frac{(\hat{\beta}_1-\beta)^2}{\dot{\Phi}_1}\ddot{\Phi}_2\Pi^\top\Pi/\sqrt{K}+o_P(1)=o_P(1).$$

If Assumptions 1-3 hold and $\Pi^\top\Pi/\sqrt{K}=O(1)$, then by results in Steps 1 and 3, we have

$$\begin{aligned}(\hat{\beta}-\beta)^2\Pi^\top\Pi/\sqrt{K} &\leqslant \left(\frac{\ddot{\Phi}_2^{1/2}}{\dot{\Phi}_1^{1/2}+\ddot{\Phi}_2^{1/2}}\right)^2\times(\hat{\beta}_1-\beta)^2\Pi^\top\Pi/\sqrt{K}\\ &\quad+\left(\frac{\dot{\Phi}_1^{1/2}}{\dot{\Phi}_1^{1/2}+\ddot{\Phi}_2^{1/2}}\right)^2\times(\hat{\beta}_2-\beta)^2\Pi^\top\Pi/\sqrt{K}\\ &=(\hat{\beta}_1-\beta)^2\times O_P(1)+\dot{\Phi}_1\times O_P(1)=o_P(1).\end{aligned}$$

If Assumptions 1, $2'$ and 3 hold and $\Pi^\top\Pi/\sqrt{K}\to\infty$, by results in Steps 2 and 4, we have

$$\begin{aligned}(\hat{\beta}-\beta)^2\Pi^\top\Pi/\sqrt{K} &\leqslant \left(\frac{\ddot{\Phi}_2^{1/2}}{\dot{\Phi}_1^{1/2}+\ddot{\Phi}_2^{1/2}}\right)^2\times(\hat{\beta}_1-\beta)^2\Pi^\top\Pi/\sqrt{K}\\ &\quad+\left(\frac{\dot{\Phi}_1^{1/2}}{\dot{\Phi}_1^{1/2}+\ddot{\Phi}_2^{1/2}}\right)^2\times(\hat{\beta}_2-\beta)^2\Pi^\top\Pi/\sqrt{K}\\ &\leqslant \frac{(\hat{\beta}_1-\beta)^2}{\dot{\Phi}_1}\ddot{\Phi}_2\Pi^\top\Pi/\sqrt{K}+o_P(1)=o_P(1).\end{aligned}$$

Therefore, we have established the desired results. Next, we focus on proving results in Steps 1–4.

Step 1: Assumptions 1-3 hold

We have

$$\hat{\beta}_1-\beta=\frac{X^\top z\hat{A}_n z^\top e}{X^\top z\hat{A}_n z^\top X}.$$

By (C.17) and (C.18), we have

$$\frac{X^\top z\hat{A}_n z^\top X}{\Pi^\top z A_n z^\top \Pi}=\frac{(\frac{1}{r_n}X^\top z)(\frac{1}{\lambda_n}\hat{A}_n)(\frac{1}{r_n}z^\top X)}{(\frac{1}{r_n}\Pi^\top z)(\frac{1}{\lambda_n}A_n)(\frac{1}{r_n}z^\top \Pi)}=1+o_P(1). \tag{C.20}$$

In addition, we have

$$\begin{aligned}
\frac{X^\top z\hat{A}_n z^\top e}{\sqrt{\Pi^\top z A_n \Omega A_n z^\top \Pi}} &= \frac{(\frac{1}{r_n}X^\top z)(\frac{1}{\lambda_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top e)}{\sqrt{(\frac{1}{r_n}\Pi^\top z)(\frac{1}{\lambda_n}A_n)(\frac{1}{n}\Omega)(\frac{1}{\lambda_n}A_n)(\frac{1}{r_n}z^\top\Pi)}} \\
&= \frac{(\frac{1}{r_n}\Pi^\top z)(\frac{1}{\lambda_n}A_n)(\frac{1}{\sqrt{n}}z^\top \tilde{e}) + o_P(1)}{\sqrt{(\frac{1}{r_n}\Pi^\top z)(\frac{1}{\lambda_n}A_n)(\frac{1}{n}\Omega)(\frac{1}{\lambda_n}A_n)(\frac{1}{r_n}z^\top\Pi)}} \\
&= \frac{\Pi^\top z A_n z^\top \tilde{e}}{\sqrt{\Pi^\top z A_n \Omega A_n z^\top \Pi}} + o_p(1),
\end{aligned}$$

where the last equality holds because by Assumptions 1 and 2,

$$\left(\frac{1}{r_n}\Pi^\top z\right)\left(\frac{1}{\lambda_n}A_n\right)\left(\frac{1}{n}\Omega\right)\left(\frac{1}{\lambda_n}A_n\right)\left(\frac{1}{r_n}z^\top\Pi\right) \geqslant \lambda_{\min}\left(\left(\frac{1}{\lambda_n}A_n\right)\left(\frac{1}{n}\Omega\right)\left(\frac{1}{\lambda_n}A_n\right)\right) \geqslant \frac{1}{C}.$$

Combining the above results and recalling

$$\Phi_1 = (\Pi^\top z A_n z^\top \Pi)^{-1}(\Pi^\top z A_n \Omega A_n z^\top \Pi)(\Pi^\top z A_n z^\top \Pi)^{-1},$$

we have

$$\frac{\hat{\beta}_1 - \beta}{\sqrt{\Phi_1}} = \frac{\Pi^\top z A_n z^\top \Pi}{X^\top z \hat{A}_n z^\top X} \times \frac{X^\top z\hat{A}_n z^\top e}{\sqrt{\Pi^\top z A_n \Omega A_n z^\top \Pi}} = O_P(1),$$

and

$$\Phi_1 = \frac{\Pi^\top z A_n \Omega A_n z^\top \Pi}{(\Pi^\top z A_n z^\top \Pi)^2} = O\left(\frac{n}{r_n^2}\right) = o(1), \tag{C.21}$$

which further imply

$$\hat{\beta}_1 - \beta = O_P(\sqrt{\Phi_1}) = o_P(1).$$

Given the consistency of $\hat{\beta}_1$, we can apply Lemma B.7 to show that $\dot{\Phi}_1/\Phi_1 \xrightarrow{p} 1$, which further implies that

$$\dot{\Phi}_1 \xrightarrow{p} 0 \quad \text{and} \quad \frac{(\hat{\beta}_1 - \beta)^2}{\dot{\Phi}_1} = O_P(1).$$

Step 2: Assumptions 1 and 3 hold, and $\Pi^\top \Pi / \sqrt{K} \to \infty$

We have

$$\hat{\beta}_2 - \beta = \frac{X^\top (P - \bar{P}) e}{X^\top (P - \bar{P}) X},$$

and

$$X^\top (P - \bar{P}) e = \hat{\Pi}^\top \tilde{e} + \tilde{V}^\top (P - \bar{P}) \tilde{e} + \tilde{V}^\top P_W \bar{P} \tilde{e} + \tilde{V}^\top \bar{P} P_W \tilde{e} - \tilde{V}^\top P_W \bar{P} P_W \tilde{e}, \tag{C.22}$$

where we use the fact that $P_W P = P P_W = 0_{n \times n}$ since $Z = M_W \tilde{Z}$. For the third term on the RHS of (C.22), as $\Sigma \geqslant C(\Pi^\top \Pi + K)$ by Lemma B.5, we have

$$\tilde{V}^\top P_W \bar{P} \tilde{e} = O_P(1) = o_P(\sqrt{\Sigma})$$

by Lemma B.3. Following the similar argument, we can show that

$$\tilde{V}^\top \bar{P} P_W \tilde{e} = o_P(\sqrt{\Sigma}) \quad \text{and} \quad \tilde{V}^\top P_W \bar{P} P_W \tilde{e} = o_P(\sqrt{\Sigma}).$$

Then, by (C.22), we have

$$X^\top (P - \bar{P}) e = \hat{\Pi}^\top \tilde{e} + \tilde{V}^\top (P - \bar{P}) \tilde{e} + o_P(\sqrt{\Sigma}). \tag{C.23}$$

In addition, we have

$$X^\top (P - \bar{P}) X = \Pi^\top (P - \bar{P}) \Pi + 2 \hat{\Pi}^\top \tilde{V} + \tilde{V}^\top M_W (P - \bar{P}) M_W \tilde{V}$$

$$= \Pi^\top(P-\bar{P})\Pi + 2\hat{\Pi}^\top\tilde{V} + \tilde{V}^\top(P-\bar{P})\tilde{V} + 2\tilde{V}^\top P_W(P-\bar{P})\tilde{V} + \tilde{V}^\top P_W(P-\bar{P})P_W\tilde{V}$$

$$= \Pi^\top(P-\bar{P})\Pi + 2\hat{\Pi}^\top\tilde{V} + \tilde{V}^\top(P-\bar{P})\tilde{V} - 2\tilde{V}^\top P_W\bar{P}\tilde{V} - \tilde{V}^\top P_W\bar{P}P_W\tilde{V}$$

$$= \Pi^\top(P-\bar{P})\Pi + 2\hat{\Pi}^\top\tilde{V} + \tilde{V}^\top(P-\bar{P})\tilde{V} + o_P(\sqrt{\Sigma})$$

$$= \Pi^\top(P-\bar{P})\Pi + O_P(\sqrt{\Sigma}), \tag{C.24}$$

where the second last equality is by

$$\tilde{V}^\top P_W\bar{P}\tilde{V} = o_P(\sqrt{\Sigma}) \quad \text{and} \quad \tilde{V}^\top P_W\bar{P}P_W\tilde{V} = o_P(\sqrt{\Sigma})$$

due to the same argument above, and the last equality holds because

$$\mathbb{V}\left(\hat{\Pi}^\top\tilde{V}\right) = O(\Pi^\top\Pi) = O(\Sigma)$$

and

$$\mathbb{V}\left(\tilde{V}^\top(P-\bar{P})\tilde{V}\right) \leqslant C||P-\bar{P}||_F^2 \leqslant CK = O(\Sigma).$$

Combining (C.23) and (C.24), we have

$$\hat{\beta}_2 - \beta = \frac{X^\top(P-\bar{P})e}{X^\top(P-\bar{P})X} = \frac{\hat{\Pi}^\top\tilde{e} + \tilde{V}^\top(P-\bar{P})\tilde{e} + o_P(\sqrt{\Sigma})}{\Pi^\top(P-\bar{P})\Pi + O_P(\sqrt{\Sigma})}.$$

In addition, we have

$$\Phi_2 = \frac{\Sigma}{\left(\Pi^\top(P-\bar{P})\Pi\right)^2} = O\left(\frac{\Pi^\top\Pi + K}{\left(\Pi^\top\Pi\right)^2}\right) = o(1) \tag{C.25}$$

because as $\Pi^\top\Pi/\sqrt{K} \to \infty$, we have $\Pi^\top\Pi \to \infty$; this also implies that

$$\frac{X^\top(P-\bar{P})X}{\Pi^\top(P-\bar{P})\Pi} = \frac{\Pi^\top(P-\bar{P})\Pi + O_P(\sqrt{\Sigma})}{\Pi^\top(P-\bar{P})\Pi} = 1 + o_P(1). \tag{C.26}$$

Therefore, we have

$$\frac{|\hat{\beta}_2 - \beta|}{\sqrt{\Phi_2}} = \frac{\left(\hat{\Pi}^\top \tilde{e} + \tilde{V}^\top (P - \bar{P})\tilde{e}\right)/\sqrt{\Sigma} + o_P(1)}{1 + O_P\left(\sqrt{\Sigma}/|\Pi^\top (P - \bar{P})\Pi|\right)} = O_P(1),$$

where we use the fact that

$$\hat{\Pi}^\top \tilde{e} = O_P(\sqrt{\Pi^\top \Pi}) = O_P(\sqrt{\Sigma})$$

and

$$\tilde{V}^\top (P - \bar{P})\tilde{e} = O_P(\sqrt{K}) = O_P(\sqrt{\Sigma}).$$

This also implies

$$\begin{aligned}
(\hat{\beta}_2 - \beta)^2 \Pi^\top \Pi/\sqrt{K} &= O_P(\Phi_2 \Pi^\top \Pi/\sqrt{K}) \\
&= O_P\left(\frac{\Pi^\top \Pi + K}{(\Pi^\top \Pi)^2} \times \frac{\Pi^\top \Pi}{\sqrt{K}}\right) \\
&= O_P\left(\frac{1}{\sqrt{K}} + \frac{\sqrt{K}}{\Pi^\top \Pi}\right) = o_P(1).
\end{aligned}$$

Given the consistency of $\hat{\beta}_2$, we can apply Lemma B.7 to show that $\ddot{\Phi}_2/\Phi_2 \xrightarrow{p} 1$, which further implies that

$$\ddot{\Phi}_2 \Pi^\top \Pi/\sqrt{K} = o_P(1).$$

Step 3: Assumptions 1 and 3 hold, and $\Pi^\top \Pi/\sqrt{K}$ is bounded

Note that $\Pi^\top (P - \bar{P})\Pi/\sqrt{K}$ is bounded in this case. In addition, let

$$\Gamma_{\tilde{V},\tilde{V}} = \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left[\left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{V}_{[h]}\right)^2 + \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{V}_{[h]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]} \tilde{V}_{[g]}\right)\right],$$

$$\Gamma_{\tilde{V},\tilde{e}} = \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left[\left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right)^2 + \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]} \tilde{e}_{[g]}\right)\right],$$

$$\Lambda = \sum_{g,h \in [G]^2, g \neq h} \mathbb{E}\left[\left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right)\left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{V}_{[h]}\right) + \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]} \tilde{V}_{[g]}\right)\right],$$

then by Assumption 3.4 we have $1/C \leqslant \Gamma_{\tilde{V},\tilde{V}}/K \leqslant C$, $1/C \leqslant \Gamma_{\tilde{V},\tilde{e}}/K \leqslant C$ and $|\Lambda|/\sqrt{\Gamma_{\tilde{V},\tilde{V}}\Gamma_{\tilde{V},\tilde{e}}} \leqslant C < 1$. We shall argue along the subsequence where $\Pi^\top(P - \bar{P})\Pi/\sqrt{K} \to \gamma \in \Re$ and

$$\frac{1}{K}\begin{pmatrix} \Gamma_{\tilde{V},\tilde{V}} & \Lambda \\ \Lambda & \Gamma_{\tilde{V},\tilde{e}} \end{pmatrix} \to \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \equiv \Gamma,$$

where $\Gamma > 0$ (in the matrix sense) by $\Gamma_{\tilde{V},\tilde{V}}/K \geqslant 1/C > 0$, $\Gamma_{\tilde{V},\tilde{e}}/K \geqslant 1/C > 0$ and $|\Lambda/K|/\sqrt{(\Gamma_{\tilde{V},\tilde{V}}/K)(\Gamma_{\tilde{V},\tilde{e}}/K)} \leqslant C < 1$.

By Assumption 3, we have

$$\begin{pmatrix} \frac{1}{\sqrt{K}}\tilde{V}^\top(P - \bar{P})\tilde{V} \\ \frac{1}{\sqrt{K}}\tilde{V}^\top(P - \bar{P})\tilde{e} \end{pmatrix} \rightsquigarrow \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}\right),$$

which can be proved by following the same steps as in the proof of Lemma B.10.

In addition, by (C.24), we have

$$\begin{aligned}\frac{1}{\sqrt{K}}X^\top(P - \bar{P})X &= \frac{1}{\sqrt{K}}\Pi^\top(P - \bar{P})\Pi + \frac{2}{\sqrt{K}}\hat{\Pi}^\top\tilde{V} \\ &\quad + \frac{1}{\sqrt{K}}\tilde{V}^\top(P - \bar{P})\tilde{V} - \frac{2}{\sqrt{K}}\tilde{V}^\top P_W \bar{P}\tilde{V} - \frac{1}{\sqrt{K}}\tilde{V}^\top P_W \bar{P} P_W \tilde{V} \\ &\rightsquigarrow \mathcal{N}(\gamma, \Gamma_{11}),\end{aligned}$$

where we use the facts that

$$\mathbb{V}\left(\frac{1}{\sqrt{K}}\hat{\Pi}\tilde{V}\right) = O\left(\frac{\Pi^\top\Pi}{K}\right) = o(1),$$

$$\frac{2}{\sqrt{K}}\tilde{V}^\top P_W\bar{P}\tilde{V} = O_P\left(1/\sqrt{K}\right) = o_P(1), \quad \text{and} \quad \frac{1}{\sqrt{K}}\tilde{V}^\top P_W\bar{P}P_W\tilde{V} = O_P\left(1/\sqrt{K}\right) = o_P(1).$$

This implies

$$\frac{1}{\frac{1}{\sqrt{K}}X^\top(P-\bar{P})X} = O_P(1). \tag{C.27}$$

Similarly, by (C.22), we have

$$\begin{aligned}\frac{1}{\sqrt{K}}X^\top(P-\bar{P})e &= \frac{1}{\sqrt{K}}\hat{\Pi}^\top\tilde{e} + \frac{1}{\sqrt{K}}\tilde{V}^\top(P-\bar{P})\tilde{e}\\ &\quad + \frac{1}{\sqrt{K}}\tilde{V}^\top P_W\bar{P}\tilde{e} + \frac{1}{\sqrt{K}}\tilde{V}^\top\bar{P}P_W\tilde{e} - \frac{1}{\sqrt{K}}\tilde{V}^\top P_W\bar{P}P_W\tilde{e}\\ &= \frac{1}{\sqrt{K}}\tilde{V}^\top(P-\bar{P})\tilde{e} + o_P(1) = O_P(1),\end{aligned}$$

which further implies

$$\hat{\Delta}_2 = \frac{X^\top(P-\bar{P})e}{X^\top(P-\bar{P})X} = \frac{\frac{1}{\sqrt{K}}X^\top(P-\bar{P})e}{\frac{1}{\sqrt{K}}X^\top(P-\bar{P})X} = O_P(1).$$

Next, we analyze $\ddot{\Phi}_2$. Note that

$$\frac{1}{K}\ddot{\Sigma} = \frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\ddot{e}_{[g]}\right)^2 + \frac{1}{K}\sum_{g,h\in[G]^2,g\neq h}\left(\tilde{X}_{[g]}^\top Q_{[g,h]}\ddot{e}_{[h]}\right)\left(\tilde{X}_{[h]}^\top Q_{[h,g]}\ddot{e}_{[g]}\right), \tag{C.28}$$

where $\ddot{e} = Y - X\hat{\beta}_2$. For the first term on the RHS of (C.28), we have

$$\begin{aligned}&\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\ddot{e}_{[g]}\right)^2\\ &= \underbrace{\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2}_{R_{21}}\end{aligned}$$

$$+\underbrace{\frac{1}{K}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}e_{[g]}\right)^2-\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2\right)}_{R_{22}}$$

$$+\underbrace{\frac{1}{K}\left(\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\ddot{e}_{[g]}\right)^2-\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}e_{[g]}\right)^2\right)}_{R_{23}}.$$

By the fact that $\Pi^\top\Pi/K = o(1)$, we have

$$1/C \leqslant \Sigma/K \leqslant C$$

for some constant $C \in (0,\infty)$, so that we can apply Lemma B.5 to obtain

$$\begin{aligned}R_{21}&=\frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1),\\&=\frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2+o_P(1),\end{aligned}$$

and $R_{22}=o_P(1)$. For R_{23}, we have

$$\begin{aligned}R_{23}=\hat{\Delta}_2^2\times&\underbrace{\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}X_{[g]}\right)^2}_{R_{23,1}}\\&-2\times\hat{\Delta}_2\times\underbrace{\frac{1}{K}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{X}_{[h]}^\top Q_{[h,g]}X_{[g]}\right)\left(\sum_{k\in[G],k\neq g}\tilde{X}_{[k]}^\top Q_{[k,g]}e_{[g]}\right)}_{R_{23,2}}.\end{aligned}$$

By Lemma B.5 and the fact that $\Pi^\top\Pi/K = o(1)$, we have

$$R_{23,1}=\frac{1}{K}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)^2+o_P(1),$$

$$R_{23,2} = \frac{1}{K} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right) + o_P(1).$$

It follows that

$$\begin{aligned}
&\frac{1}{K} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{X}_{[h]}^\top Q_{[h,g]} \ddot{e}_{[g]} \right)^2 \\
&= \frac{1}{K} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)^2 \\
&+ \frac{\hat{\Delta}_2^2}{K} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right)^2 \\
&- \frac{2\hat{\Delta}_2}{K} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) \left(\sum_{k \in [G], k \neq g} \tilde{V}_{[k]}^\top Q_{[k,g]} \tilde{e}_{[g]} \right) + o_P(1),
\end{aligned} \tag{C.29}$$

since $\hat{\Delta}_2 = O_P(1)$. For the second term on the RHS of (C.28), we have

$$\begin{aligned}
&\frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \ddot{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \ddot{e}_{[g]} \right) \\
&= \underbrace{\frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right)}_{R_{24}} \\
&+ \underbrace{\frac{1}{K} \begin{pmatrix} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} e_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} e_{[g]} \right) \\ - \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) \end{pmatrix}}_{R_{25}} \\
&+ \underbrace{\frac{1}{K} \begin{pmatrix} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \ddot{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \ddot{e}_{[g]} \right) \\ - \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} e_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} e_{[g]} \right) \end{pmatrix}}_{R_{26}}.
\end{aligned}$$

By Lemma B.6 and the fact that $\Pi^\top \Pi / K = o(1)$, we have

$$R_{24} = \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) + o_P(1),$$

and $R_{25} = o_P(1)$. For R_{26}, we have

$$\begin{aligned} R_{26} = \hat{\Delta}^2 \times & \underbrace{\frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} X_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} X_{[g]} \right)}_{R_{26,1}} \\ & - 2 \times \hat{\Delta} \times \underbrace{\frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} X_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} e_{[g]} \right)}_{R_{26,2}}. \end{aligned}$$

By Lemma B.6 and the fact that $\Pi^\top \Pi / K = o(1)$, we have

$$\begin{aligned} R_{26,1} &= \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) + o_P(1) \\ &= \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{V}_{[g]} \right) + o_P(1), \end{aligned}$$

and

$$\begin{aligned} R_{26,2} &= \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) + o_P(1) \\ &= \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \tilde{V}_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) + o_P(1). \end{aligned}$$

It follows that

$$\begin{aligned} & \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{X}_{[g]}^\top Q_{[g,h]} \ddot{e}_{[h]} \right) \left(\tilde{X}_{[h]}^\top Q_{[h,g]} \ddot{e}_{[g]} \right) \\ = & \frac{1}{K} \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top Q_{[g,h]} \tilde{e}_{[h]} \right) \left(\tilde{V}_{[h]}^\top Q_{[h,g]} \tilde{e}_{[g]} \right) \end{aligned}$$

$$+\frac{\hat{\Delta}_2^2}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)$$
$$-\frac{2\hat{\Delta}_2}{K}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)+o_P(1). \tag{C.30}$$

since $\hat{\Delta}_2 = O_P(1)$. Combining (C.28)–(C.30), we obtain that

$$\begin{aligned}
\frac{1}{K}\ddot{\Sigma} &= \frac{\hat{\Delta}_2^2}{K}\left(\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)^2+\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)\right)\\
&\quad+\frac{1}{K}\left(\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)^2+\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\right)\\
&\quad-\frac{2\hat{\Delta}_2}{K}\Bigg[\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{V}_{[g]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\\
&\qquad\qquad+\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top Q_{[g,h]}\tilde{V}_{[h]}\right)\left(\tilde{V}_{[h]}^\top Q_{[h,g]}\tilde{e}_{[g]}\right)\Bigg]+o_P(1)\\
&= \frac{\hat{\Delta}_2^2}{K}\mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{V}\right)^2-\frac{2\hat{\Delta}_2}{K}\mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{V}\right)\left(\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right)\\
&\quad+\frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right)^2+o_P(1).
\end{aligned}$$

In addition, by Lemma B.4, we have

$$\begin{aligned}
\frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{V}\right)^2 &= \frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(P-\bar{P})\tilde{V}\right)^2+o(1)=\Gamma_{11}+o(1),\\
\frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right)^2 &= \frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(P-\bar{P})\tilde{e}\right)^2+o(1)=\Gamma_{22}+o(1),\\
\frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(Q-\bar{Q})\tilde{V}\right)\left(\tilde{V}^\top(Q-\bar{Q})\tilde{e}\right) &= \frac{1}{K}\mathbb{E}\left(\tilde{V}^\top(P-\bar{P})\tilde{V}\right)\left(\tilde{V}^\top(P-\bar{P})\tilde{e}\right)+o(1)=\Gamma_{12}+o(1).
\end{aligned}$$

Combining the above results, we have

$$\frac{1}{K}\ddot{\Sigma}=\hat{\Delta}_2^2\Gamma_{11}-2\hat{\Delta}_2\Gamma_{12}+\Gamma_{22}+o_P(1). \tag{C.31}$$

It follows that

$$
\begin{aligned}
\ddot{\Phi}_2 &= \left(\frac{1}{\sqrt{K}} X^\top (P-\bar{P}) X\right)^{-2} \left(\frac{1}{K}\ddot{\Sigma}\right) \\
&= \frac{(\frac{1}{\sqrt{K}} X^\top (P-\bar{P}) e)^2}{(\frac{1}{\sqrt{K}} X^\top (P-\bar{P}) X)^4} \times \Gamma_{11} - 2\frac{(\frac{1}{\sqrt{K}} X^\top (P-\bar{P}) e)}{(\frac{1}{\sqrt{K}} X^\top (P-\bar{P}) X)^3} \times \Gamma_{12} \\
&+ \frac{1}{(\frac{1}{\sqrt{K}} X^\top (P-\bar{P}) X)^2} \times \Gamma_{22} + o_P(1) \\
&\rightsquigarrow \bar{\Phi}_2,
\end{aligned}
$$

where the second equality is by (C.27) and (C.31), and $\bar{\Phi}_2$ is defined as

$$
\begin{aligned}
\bar{\Phi}_2 &= \frac{\eta_{\tilde{V},\tilde{e}}^2}{(\eta_{\tilde{V},\tilde{V}}+\gamma)^4}\Gamma_{11} - \frac{2\eta_{\tilde{V},\tilde{e}}}{(\eta_{\tilde{V},\tilde{V}}+\gamma)^3}\Gamma_{12} + \frac{1}{(\eta_{\tilde{V},\tilde{V}}+\gamma)^2}\Gamma_{22} \\
&= \frac{1}{(\eta_{\tilde{V},\tilde{V}}+\gamma)^4}\left(\eta_{\tilde{V},\tilde{e}}^2\Gamma_{11} - 2\eta_{\tilde{V},\tilde{e}}(\eta_{\tilde{V},\tilde{V}}+\gamma)\Gamma_{12} + (\eta_{\tilde{V},\tilde{V}}+\gamma)^2\Gamma_{22}\right) \\
&= \frac{\begin{pmatrix} -\eta_{\tilde{V},\tilde{e}} \\ \eta_{\tilde{V},\tilde{V}}+\gamma \end{pmatrix}^\top \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \begin{pmatrix} -\eta_{\tilde{V},\tilde{e}} \\ \eta_{\tilde{V},\tilde{V}}+\gamma \end{pmatrix}}{(\eta_{\tilde{V},\tilde{V}}+\gamma)^4},
\end{aligned}
$$

with

$$
\begin{pmatrix} \eta_{\tilde{V},\tilde{V}} \\ \eta_{\tilde{V},\tilde{e}} \end{pmatrix} \stackrel{d}{=} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}\right).
$$

Note that

$$
\begin{pmatrix} -\eta_{\tilde{V},\tilde{e}} \\ \eta_{\tilde{V},\tilde{V}}+\gamma \end{pmatrix}^\top \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \begin{pmatrix} -\eta_{\tilde{V},\tilde{e}} \\ \eta_{\tilde{V},\tilde{V}}+\gamma \end{pmatrix} \geqslant 0,
$$

and the equality holds if and only if $\eta_{\tilde{V},\tilde{e}} = 0$ and $\eta_{\tilde{V},\tilde{V}} + \gamma = 0$, which has probability zero since $\eta_{\tilde{V},\tilde{e}}$ has a non-degenerate normal distribution and similarly for $\eta_{\tilde{V},\tilde{V}}$. In addition, the denominator of $\bar{\Phi}_2$ is positive with probability one. Therefore, we have $\bar{\Phi}_2 > 0$ with probability one, which implies that

$$1/\ddot{\Phi}_2^{1/2} = O_P(1).$$

Step 4: Assumptions 1, 2′ and 3 hold

By Assumptions 2′.1 and 2′.2, we can argue along the subsequence where $z^\top \Pi/\sqrt{n} \to \pi \in \Re^{d_z}$,

$$\frac{1}{n}\begin{pmatrix} \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{e}_{[g]}\right)\left(z_{[g]}^\top \tilde{e}_{[g]}\right)^\top & \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{e}_{[g]}\right)\left(z_{[g]}^\top \tilde{V}_{[g]}\right)^\top \\ \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{V}_{[g]}\right)\left(z_{[g]}^\top \tilde{e}_{[g]}\right)^\top & \sum_{g\in[G]} \mathbb{E}\left(z_{[g]}^\top \tilde{V}_{[g]}\right)\left(z_{[g]}^\top \tilde{V}_{[g]}\right)^\top \end{pmatrix} \to \begin{pmatrix} \Omega_z^{\tilde{e},\tilde{e}} & \Omega_z^{\tilde{e},\tilde{V}} \\ \Omega_z^{\tilde{V},\tilde{e}} & \Omega_z^{\tilde{V},\tilde{V}} \end{pmatrix} > 0,$$

and

$$\frac{1}{n}\sum_{g\in[G]} \left(z_{[g]}^\top \Pi_{[g]}\right)\left(z_{[g]}^\top \Pi_{[g]}\right)^\top \to \Omega_z^{\Pi,\Pi} \geqslant 0,$$

in the matrix sense. In addition, note that since Assumption 2.1 holds, A_n/λ_n has eigenvalues bounded and bounded away from zero, where $\lambda_n = \lambda_{\max}(A_n)$. Therefore, without loss of generality, we assume that $A_n/\lambda_n \to A$ for some non-random positive definite matrix A with eigenvalues bounded and bounded away from zero (otherwise we argue along a further subsequence).

By Assumption 2′, we have

$$\begin{pmatrix} \frac{1}{\sqrt{n}} z^\top \tilde{e} \\ \frac{1}{\sqrt{n}} z^\top \tilde{V} \end{pmatrix} \rightsquigarrow \mathcal{N}\left(0, \begin{pmatrix} \Omega_z^{\tilde{e},\tilde{e}} & \Omega_z^{\tilde{e},\tilde{V}} \\ \Omega_z^{\tilde{V},\tilde{e}} & \Omega_z^{\tilde{V},\tilde{V}} \end{pmatrix}\right),$$

which can be proved by following the same steps as in the proof of Lemma B.10 (see also Hansen and Lee (2019) and Djogbenou, MacKinnon, and Nielsen (2019)). This also implies that

$$\frac{1}{\sqrt{n}} z^\top X \rightsquigarrow \mathcal{N}\left(\pi, \Omega_z^{\tilde{V},\tilde{V}}\right).$$

We have

$$\hat{\Delta}_1 = \frac{X^\top z \hat{A}_n z^\top e}{X^\top z \hat{A}_n z^\top X} = \frac{(\frac{1}{\sqrt{n}} X^\top z)(\frac{1}{\tilde{\lambda}_n} \hat{A}_n)(\frac{1}{\sqrt{n}} z^\top \tilde{e})}{(\frac{1}{\sqrt{n}} X^\top z)(\frac{1}{\tilde{\lambda}_n} \hat{A}_n)(\frac{1}{\sqrt{n}} z^\top X)}$$

and

$$(\frac{1}{\sqrt{n}} X^\top z)(\frac{1}{\tilde{\lambda}_n} \hat{A}_n)(\frac{1}{\sqrt{n}} z^\top \tilde{e}) = O_P(1).$$

We also have

$$\frac{1}{(\frac{1}{\sqrt{n}} X^\top z)(\frac{1}{\tilde{\lambda}_n} \hat{A}_n)(\frac{1}{\sqrt{n}} z^\top X)} \leqslant \frac{1}{\xi_n^2},$$

where ξ_n is the first element of $(\hat{A}_n/\tilde{\lambda}_n)^{1/2}(z^\top X/\sqrt{n})$, and note that $\xi_n \rightsquigarrow \xi$ where ξ has a non-degenerate normal distribution. This implies that

$$\frac{1}{(\frac{1}{\sqrt{n}} X^\top z)(\frac{1}{\tilde{\lambda}_n} \hat{A}_n)(\frac{1}{\sqrt{n}} z^\top X)} = O_P(1),$$

whence $\hat{\Delta}_1 = O_P(1)$.

Next, we analyze $\dot{\Phi}_1$. Note that

$$\begin{aligned}\frac{1}{n}\dot{\Omega} &= \left(\frac{1}{n}\dot{\Omega} - \frac{1}{n}\bar{\Omega}\right) + \left(\frac{1}{n}\bar{\Omega} - \frac{1}{n}\tilde{\Omega}\right) + \left(\frac{1}{n}\tilde{\Omega} - \frac{1}{n}\Omega\right) + \frac{1}{n}\Omega \\ &= \left(\frac{1}{n}\dot{\Omega} - \frac{1}{n}\bar{\Omega}\right) + \Omega_z^{\tilde{e},\tilde{e}} + o_P(1),\end{aligned}$$

as in the proof of Lemma B.7, where $\dot{\Omega}$ is just $\acute{\Omega}$ with $\acute{\beta} = \hat{\beta}_1$. We have

$$\frac{1}{n}\tilde{\Omega} - \frac{1}{n}\bar{\Omega} = \hat{\Delta}_1^2 \times \underbrace{\frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top X_{[g]})(X_{[g]}^\top z_{[g]})}_{R_{14}} - \hat{\Delta}_1 \times \underbrace{\frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top X_{[g]})e_{[g]}^\top z_{[g]}}_{R_{15}} - \hat{\Delta}_1 \times \underbrace{\frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top e_{[g]})(X_{[g]}^\top z_{[g]})}_{R_{16}}.$$

By using the same argument as in the proof of Lemma B.7, we have

$$\begin{aligned}
R_{14} &= \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top \Pi_{[g]})(\Pi_{[g]}^\top z_{[g]}) + \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top V_{[g]})(V_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top \Pi_{[g]})(\Pi_{[g]}^\top z_{[g]}) + \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top \tilde{V}_{[g]})(\tilde{V}_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top \Pi_{[g]})(\Pi_{[g]}^\top z_{[g]}) + \frac{1}{n}\sum_{g\in[G]}\mathbb{E}(z_{[g]}^\top \tilde{V}_{[g]})(\tilde{V}_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \Omega_z^{\Pi,\Pi} + \Omega_z^{\tilde{V},\tilde{V}} + o_P(1),
\end{aligned}$$

$$\begin{aligned}
R_{15} &= \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top V_{[g]})(e_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top \tilde{V}_{[g]})(\tilde{e}_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \frac{1}{n}\sum_{g\in[G]}\mathbb{E}(z_{[g]}^\top \tilde{V}_{[g]})(\tilde{e}_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \Omega_z^{\tilde{V},\tilde{e}} + o_P(1),
\end{aligned}$$

and

$$
\begin{aligned}
R_{16} &= \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top e_{[g]})(V_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \frac{1}{n}\sum_{g\in[G]}(z_{[g]}^\top \tilde{e}_{[g]})(\tilde{V}_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \frac{1}{n}\sum_{g\in[G]}\mathbb{E}(z_{[g]}^\top \tilde{e}_{[g]})(\tilde{V}_{[g]}^\top z_{[g]}) + o_P(1) \\
&= \Omega_z^{\tilde{e},\tilde{V}} + o_P(1).
\end{aligned}
$$

We thus obtain

$$
\frac{1}{n}\tilde{\Omega} = \hat{\Delta}_1^2\Omega_z^{\Pi,\Pi} + \hat{\Delta}_1^2\Omega_z^{\tilde{V},\tilde{V}} - \hat{\Delta}_1\Omega_z^{\tilde{V},\tilde{e}} - \hat{\Delta}_1\Omega_z^{\tilde{e},\tilde{V}} + \Omega_z^{\tilde{e},\tilde{e}} + o_P(1),
$$

since $\hat{\Delta}_1 = O_P(1)$.

Combining the above results, we have

$$
\begin{aligned}
\dot{\Phi}_1 &= \left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right)^{-2}\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{n}\tilde{\Omega})(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right) \\
&= \frac{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top \tilde{e})\right)^2}{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right)^4} \times \left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)\Omega_z^{\Pi,\Pi}(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right) \\
&+ \frac{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top \tilde{e})\right)^2}{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right)^4} \times \left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)\Omega_z^{\tilde{V},\tilde{V}}(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right) \\
&- \frac{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top \tilde{e})\right)}{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right)^3} \times \left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)\Omega_z^{\tilde{V},\tilde{e}}(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right) \\
&- \frac{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top \tilde{e})\right)}{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right)^3} \times \left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)\Omega_z^{\tilde{e},\tilde{V}}(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right) \\
&+ \frac{1}{\left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right)^2} \times \left((\frac{1}{\sqrt{n}}X^\top z)(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)\Omega_z^{\tilde{e},\tilde{e}}(\frac{1}{\tilde{\lambda}_n}\hat{A}_n)(\frac{1}{\sqrt{n}}z^\top X)\right) + o_P(1),
\end{aligned}
$$

and thus $\dot{\Phi}_1 \rightsquigarrow \bar{\Phi}_1$, by the continuous mapping theorem, where

$$\begin{aligned}
\bar{\Phi}_1 &= \frac{\left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)^2}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^4} \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\Pi,\Pi}A(\zeta_{\tilde{V}}+\pi)\right) \\
&+ \frac{\left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)^2}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^4} \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{V},\tilde{V}}A(\zeta_{\tilde{V}}+\pi)\right) \\
&- \frac{\left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^3} \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{V},\tilde{e}}A(\zeta_{\tilde{V}}+\pi)\right) \\
&- \frac{\left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^3} \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{e},\tilde{V}}A(\zeta_{\tilde{V}}+\pi)\right) \\
&+ \frac{1}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^2} \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{e},\tilde{e}}A(\zeta_{\tilde{V}}+\pi)\right) \\
&= \frac{1}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^4} \Big\{ \left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)^2 \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\Pi,\Pi}A(\zeta_{\tilde{V}}+\pi)\right) \\
&+ \left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)^2 \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{V},\tilde{V}}A(\zeta_{\tilde{V}}+\pi)\right) \\
&- \left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right) \times \left((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi)\right) \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{V},\tilde{e}}A(\zeta_{\tilde{V}}+\pi)\right) \\
&- \left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right) \times \left((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi)\right) \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{e},\tilde{V}}A(\zeta_{\tilde{V}}+\pi)\right) \\
&+ \left((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi)\right)^2 \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\tilde{e},\tilde{e}}A(\zeta_{\tilde{V}}+\pi)\right) \Big\} \\
&= \frac{\left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)^2 \times \left((\zeta_{\tilde{V}}+\pi)^\top A\Omega_z^{\Pi,\Pi}A(\zeta_{\tilde{V}}+\pi)\right)}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^4} \\
&+ \frac{\left[\begin{pmatrix} -\left((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi)\right)(A(\zeta_{\tilde{V}}+\pi)) \\ \left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)(A(\zeta_{\tilde{V}}+\pi)) \end{pmatrix}^\top \begin{pmatrix} \Omega_z^{\tilde{e},\tilde{e}} & \Omega_z^{\tilde{e},\tilde{V}} \\ \Omega_z^{\tilde{V},\tilde{e}} & \Omega_z^{\tilde{V},\tilde{V}} \end{pmatrix} \times \begin{pmatrix} -\left((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi)\right)(A(\zeta_{\tilde{V}}+\pi)) \\ \left((\zeta_{\tilde{V}}+\pi)^\top A\zeta_{\tilde{e}}\right)(A(\zeta_{\tilde{V}}+\pi)) \end{pmatrix} \right]}{((\zeta_{\tilde{V}}+\pi)^\top A(\zeta_{\tilde{V}}+\pi))^4},
\end{aligned}$$

with

$$\begin{pmatrix} \zeta_{\tilde{e}} \\ \zeta_{\tilde{V}} \end{pmatrix} \overset{d}{=} \mathcal{N}\left(0, \begin{pmatrix} \Omega_z^{\tilde{e},\tilde{e}} & \Omega_z^{\tilde{e},\tilde{V}} \\ \Omega_z^{\tilde{V},\tilde{e}} & \Omega_z^{\tilde{V},\tilde{V}} \end{pmatrix}\right).$$

Finally, we note that

$$\left((\zeta_{\tilde V}+\pi)^\top A\zeta_{\tilde e}\right)^2\times\left((\zeta_{\tilde V}+\pi)^\top A\Omega_z^{\Pi,\Pi}A(\zeta_{\tilde V}+\pi)\right)\geqslant 0,$$

and

$$\left[\begin{array}{c}\begin{pmatrix}-\left((\zeta_{\tilde V}+\pi)^\top A(\zeta_{\tilde V}+\pi)\right)\left(A(\zeta_{\tilde V}+\pi)\right)\\ \left((\zeta_{\tilde V}+\pi)^\top A\zeta_{\tilde e}\right)\left(A(\zeta_{\tilde V}+\pi)\right)\end{pmatrix}^\top\begin{pmatrix}\Omega_z^{\tilde e,\tilde e} & \Omega_z^{\tilde e,\tilde V}\\ \Omega_z^{\tilde V,\tilde e} & \Omega_z^{\tilde V,\tilde V}\end{pmatrix}\\ \times\begin{pmatrix}-\left((\zeta_{\tilde V}+\pi)^\top A(\zeta_{\tilde V}+\pi)\right)\left(A(\zeta_{\tilde V}+\pi)\right)\\ \left((\zeta_{\tilde V}+\pi)^\top A\zeta_{\tilde e}\right)\left(A(\zeta_{\tilde V}+\pi)\right)\end{pmatrix}\end{array}\right]\geqslant 0,$$

and the equalities hold if and only if

$$\left((\zeta_{\tilde V}+\pi)^\top A(\zeta_{\tilde V}+\pi)\right)\left(A(\zeta_{\tilde V}+\pi)\right)=0,$$
$$\left((\zeta_{\tilde V}+\pi)^\top A\zeta_{\tilde e}\right)\left(A(\zeta_{\tilde V}+\pi)\right)=0.$$

Given that A is positive definite, the above two equalities hold if and only if $\zeta_{\tilde V}+\pi=0$, which has probability zero since $\zeta_{\tilde V}$ has a non-degenerate normal distribution. In addition, the denominator of $\bar\Phi_1$ is positive with probability one, which implies

$$1/\dot\Phi_1^{1/2}=O_P(1).$$

This concludes the proof. □

C.9 Proof of Lemma B.9

By Lemma B.8, we have $\hat\beta\xrightarrow{p}\beta$ and $(\hat\beta-\beta)^2\Pi^\top\Pi/\sqrt{K}=o_P(1)$.

For $T(\beta_0)$, by Lemma B.7,

$$\frac{\Psi}{\widehat{\Psi}} \xrightarrow{p} 1,$$

which, by (C.20), implies that

$$\frac{\Phi_1}{\widehat{\Phi}_1} \xrightarrow{p} 1.$$

We have

$$\begin{aligned} T(\beta_0) &= \frac{(X^\top z\hat{A}_n z^\top X)^{-1}(X^\top z\hat{A}_n z^\top e)}{\sqrt{\widehat{\Phi}_1}} + a_1\delta + \left(\frac{d_n}{\sqrt{\widehat{\Phi}_1}} - a_1\right)\delta \\ &= \frac{(X^\top z\hat{A}_n z^\top X)^{-1}(X^\top z\hat{A}_n z^\top e)}{\sqrt{\widehat{\Phi}_1}} + a_1\delta + o_P(1), \end{aligned}$$

where the second equality holds because

$$\frac{d_n}{\sqrt{\widehat{\Phi}_1}} - a_1 = \frac{d_n}{\sqrt{\Phi_1}}\frac{\sqrt{\Phi_1}}{\sqrt{\widehat{\Phi}_1}} - a_1 = o_p(1).$$

In addition, we have

$$\begin{aligned} \frac{(X^\top z\hat{A}_n z^\top X)^{-1}(X^\top z\hat{A}_n z^\top e)}{\sqrt{\widehat{\Phi}_1}} &= \frac{(X^\top z\hat{A}_n z^\top X)^{-1}}{\sqrt{(X^\top z\hat{A}_n z^\top X)^{-2}}} \times \frac{(X^\top z\hat{A}_n z^\top e)}{\sqrt{\hat{\Psi}}} \\ &= \sqrt{\frac{\Psi}{\hat{\Psi}}} \times \frac{(X^\top z\hat{A}_n z^\top e)}{\sqrt{\Psi}} \times (1 + o_P(1)) \\ &= \frac{(\Pi^\top z A_n z^\top \tilde{e})}{\sqrt{\Pi^\top z A_n \Omega A_n z^\top \Pi}} \times (1 + o_P(1)) \\ &= \frac{(\Pi^\top z A_n z^\top \tilde{e})}{\sqrt{\Pi^\top z A_n \Omega A_n z^\top \Pi}} + o_P(1), \end{aligned}$$

whence

$$T(\beta_0) = \frac{1}{\sqrt{\Psi}} \sum_{g \in [G]} \acute{\Pi}_{[g]}^\top \tilde{e}_{[g]} + a_1 \delta + o_P(1).$$

For $LM(\beta_0)$, if $\Pi^\top \Pi / \sqrt{K} \to \infty$, we have

$$\begin{aligned}
LM(\beta_0) &= \frac{X^\top (P - \bar{P}) e}{\sqrt{\hat{\Sigma}}} + a_2 \delta + \left(\frac{d_n (X^\top (P - \bar{P}) X)}{\sqrt{\hat{\Sigma}}} - a_2 \right) \delta \\
&= \frac{X^\top (P - \bar{P}) e}{\sqrt{\hat{\Sigma}}} + a_2 \delta + \left(\frac{d_n}{\sqrt{\Phi_2}} \times (1 + o_P(1)) - a_2 \right) \delta \\
&= \frac{X^\top (P - \bar{P}) e}{\sqrt{\hat{\Sigma}}} + a_2 \delta + o_P(1),
\end{aligned}$$

where the second equality holds because by Lemma B.7,

$$\frac{\Sigma}{\hat{\Sigma}} \xrightarrow{p} 1,$$

which, by (C.26), implies that

$$\frac{\Phi_2 (X^\top (P - \bar{P}) X)^2}{\hat{\Sigma}} = \frac{\Phi_2 (X^\top (P - \bar{P}) X)^2}{\Sigma} \frac{\Sigma}{\hat{\Sigma}} = \frac{(X^\top (P - \bar{P}) X)^2}{(\Pi^\top (P - \bar{P}) \Pi)^2} \frac{\Sigma}{\hat{\Sigma}} \xrightarrow{p} 1.$$

Alternatively, if $\Pi^\top \Pi / \sqrt{K} = O(1)$, we have

$$\begin{aligned}
LM(\beta_0) &= \frac{X^\top (P - \bar{P}) e}{\sqrt{\hat{\Sigma}}} + \frac{d_n \delta (X^\top (P - \bar{P}) X)}{\sqrt{\hat{\Sigma}}} \\
&= \frac{X^\top (P - \bar{P}) e}{\sqrt{\hat{\Sigma}}} + o_P(1),
\end{aligned}$$

where we use the fact that

$$\frac{X^\top (P - \bar{P}) X}{\sqrt{\hat{\Sigma}}} = \frac{X^\top (P - \bar{P}) X}{\sqrt{\Sigma}} \times \frac{\Sigma}{\hat{\Sigma}} = \frac{\Pi^\top (P - \bar{P}) \Pi + O_P(\sqrt{\Sigma})}{\sqrt{\Sigma}} \times O_P(1) = O_P(1)$$

by (C.24), $\Pi^\top\Pi/\sqrt{K} = O(1)$, and

$$\frac{1}{\sqrt{\Phi_2}} = \sqrt{\frac{(\Pi^\top(P-\bar{P})\Pi)^2}{\Sigma}} = O(\frac{\Pi^\top\Pi}{\sqrt{K}}) = O(1),$$

Assumption 4 implies that $a_2 = \lim_{n\to\infty} d_n/\sqrt{\Phi_2} = 0$ in this case. In either case, we obtain

$$LM(\beta_0) = \frac{X^\top(P-\bar{P})e}{\sqrt{\hat{\Sigma}}} + a_2\delta + o_P(1).$$

In addition, we have

$$\begin{aligned}
\frac{X^\top(P-\bar{P})e}{\sqrt{\hat{\Sigma}}} &= \sqrt{\frac{\Sigma}{\hat{\Sigma}}} \times \frac{1}{\sqrt{\Sigma}}\left(\hat{\Pi}^\top\tilde{e} + \tilde{V}^\top(P-\bar{P})\tilde{e} + \tilde{V}^\top P_W\bar{P}\tilde{e} + \tilde{V}^\top\bar{P}P_W\tilde{e} - \tilde{V}^\top P_W\bar{P}P_W\tilde{e}\right) \\
&= \sqrt{\frac{\Sigma}{\hat{\Sigma}}} \times \frac{1}{\sqrt{\Sigma}}\left(\hat{\Pi}^\top\tilde{e} + \tilde{V}^\top(P-\bar{P})\tilde{e} + o_P(\sqrt{\Sigma})\right), \\
&= \frac{1}{\sqrt{\Sigma}}\left(\hat{\Pi}^\top\tilde{e} + \tilde{V}^\top(P-\bar{P})\tilde{e}\right) + o_P(1),
\end{aligned}$$

where the second equality holds by $\Sigma \geqslant C(\Pi^\top\Pi + K) \to \infty$ as shown in Lemmas B.3 and B.5, and the last equality holds by Lemma B.7. It follows that

$$LM(\beta_0) = \frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\hat{\Pi}_{[g]}^\top\tilde{e}_{[g]} + \sum_{g,h\in[G]^2, g\neq h}\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right) + a_2\delta + o_P(1).$$

For AR, we have

$$e^\top(P-\bar{P})e - \tilde{e}^\top(P-\bar{P})\tilde{e} = \tilde{e}^\top P_W\bar{P}\tilde{e} + \tilde{e}^\top\bar{P}P_W\tilde{e} - \tilde{e}^\top P_W\bar{P}P_W\tilde{e} = O_P(1)$$

by Lemma B.3, and $\Upsilon \geqslant CK$ by (C.19). This implies that

$$\frac{1}{\sqrt{\Upsilon}}\left(\sum_{g,h\in[G]^2, g\neq h} e_{[g]}^\top P_{[g,h]}e_{[h]} - \sum_{g,h\in[G]^2, g\neq h}\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right) = O_P(K^{-1/2}) = o_P(1).$$

We also have

$$\begin{aligned}
\hat{e}^\top(P-\bar{P})\hat{e} - e^\top(P-\bar{P})e &= (\hat{\beta}-\beta)^2 X^\top(P-\bar{P})X - 2(\hat{\beta}-\beta)X^\top(P-\bar{P})e \\
&= (\hat{\beta}-\beta)^2\left(\Pi^\top(P-\bar{P})\Pi + O_P(\sqrt{\Sigma})\right) \\
&\quad - 2(\hat{\beta}-\beta)\left(\tilde{\Pi}^\top\tilde{e} + \tilde{V}^\top(P-\bar{P})\tilde{e} + o_P(\sqrt{\Sigma})\right) \\
&= (\hat{\beta}-\beta)^2 O_P(\Pi^\top\Pi + \sqrt{\Sigma}) + (\hat{\beta}-\beta)O_P(\sqrt{\Sigma}) \\
&= \left((\hat{\beta}-\beta)^2\Pi^\top\Pi + (\hat{\beta}-\beta)^2\sqrt{K} + (\hat{\beta}-\beta)\sqrt{K}\right)\times O_P(1)
\end{aligned}$$

by (C.22) and (C.24), $\Pi^\top\Pi = O(\sqrt{K})$, and $\Sigma = O(\Pi^\top\Pi + K) = O(K)$. Then by $\hat{\beta} \xrightarrow{p} \beta$ and $(\hat{\beta}-\beta)^2\Pi^\top\Pi/\sqrt{K} = o_P(1)$, we have

$$\begin{aligned}
&\frac{1}{\sqrt{\Upsilon}}\left(\sum_{g,h\in[G]^2, g\neq h}\hat{e}_{[g]}^\top P_{[g,h]}\hat{e}_{[h]} - \sum_{g,h\in[G]^2, g\neq h} e_{[g]}^\top P_{[g,h]} e_{[h]}\right) \\
&= \left(\frac{(\hat{\beta}-\beta)^2\Pi^\top\Pi}{\sqrt{K}} + (\hat{\beta}-\beta)^2 + (\hat{\beta}-\beta)\right)\times O_P(1) \\
&= o_P(1).
\end{aligned}$$

It follows that, by Lemma B.7, we have

$$\begin{aligned}
AR &= \sqrt{\frac{\Upsilon}{\hat{\Upsilon}}}\times\frac{1}{\sqrt{\Upsilon}}\sum_{g,h\in[G]^2, g\neq h}\hat{e}_{[g]}^\top P_{[g,h]}\hat{e}_{[h]} \\
&= \sqrt{\frac{\Upsilon}{\hat{\Upsilon}}}\times\frac{1}{\sqrt{\Upsilon}}\left(\sum_{g,h\in[G]^2, g\neq h}\hat{e}_{[g]}^\top P_{[g,h]}\hat{e}_{[h]} - \sum_{g,h\in[G]^2, g\neq h} e_{[g]}^\top P_{[g,h]} e_{[h]}\right) \\
&+ \sqrt{\frac{\Upsilon}{\hat{\Upsilon}}}\times\frac{1}{\sqrt{\Upsilon}}\left(\sum_{g,h\in[G]^2, g\neq h} e_{[g]}^\top P_{[g,h]} e_{[h]} - \sum_{g,h\in[G]^2, g\neq h}\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right) \\
&+ \sqrt{\frac{\Upsilon}{\hat{\Upsilon}}}\times\frac{1}{\sqrt{\Upsilon}}\sum_{g,h\in[G]^2, g\neq h}\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]} \\
&= \frac{1}{\sqrt{\Upsilon}}\sum_{g,h\in[G]^2, g\neq h}\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]} + o_P(1).
\end{aligned}$$

This concludes the proof. □

C.10 Proof of Lemma B.10

Note that by Assumption 1.2, we have $G \to \infty$ as $n \to \infty$. The proof follows the same steps as in Chao et al. (2012) to check the conditions for the martingale central limit theorem; see, for example, Hall and Heyde (1980).

Step 1: Construct martingale difference array. Let

$$\begin{aligned}
\rho_{1n} &= \operatorname{cov}\left[\frac{1}{\sqrt{\Psi}}\sum_{g\in[G]}\acute{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}, \frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]} + \sum_{g,h\in[G]^2,g\neq h}\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\right] \\
&= \frac{1}{\sqrt{\Psi\Sigma}}\sum_{g\in[G]}\mathbb{E}\left[\left(\acute{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right],
\end{aligned}$$

such that $\rho_1 = \lim_{n\to\infty}\rho_{1n}$, and

$$\begin{aligned}
\rho_{2n} &= \operatorname{cov}\left[\frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]} + \sum_{g,h\in[G]^2,g\neq h}\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right), \frac{1}{\sqrt{\Upsilon}}\sum_{g,h\in[G]^2,g\neq h}\tilde{e}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right] \\
&= \frac{2}{\sqrt{\Sigma\Upsilon}}\sum_{g,h\in[G]^2,g\neq h}\mathbb{E}\left[\left(\tilde{V}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{e}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right)\right],
\end{aligned}$$

such that $\rho_2 = \lim_{n\to\infty}\rho_{2n}$, and note that

$$\operatorname{cov}\left[\frac{1}{\sqrt{\Psi}}\sum_{g\in[G]}\acute{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}, \frac{1}{\sqrt{\Upsilon}}\sum_{g,h\in[G]^2,g\neq h}\tilde{e}_{[g]}^{\top}P_{[g,h]}\tilde{e}_{[h]}\right] = 0.$$

The assumptions for ρ_1 and ρ_2 in Theorem 4.1 ensure that

$$\begin{pmatrix} 1 & \rho_{1n} & 0 \\ \rho_{1n} & 1 & \rho_{2n} \\ 0 & \rho_{2n} & 1 \end{pmatrix}^{-1}$$

exists for all n large enough, and by the Slutsky theorem, the result would follow if

$$\begin{pmatrix} 1 & \rho_{1n} & 0 \\ \rho_{1n} & 1 & \rho_{2n} \\ 0 & \rho_{2n} & 1 \end{pmatrix}^{-1/2} \begin{pmatrix} \frac{1}{\sqrt{\Psi}} \sum_{g \in [G]} \acute{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} \\ \frac{1}{\sqrt{\Sigma}} \left(\sum_{g \in [G]} \hat{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} + \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \right) \\ \frac{1}{\sqrt{\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \tilde{e}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \end{pmatrix}$$

$$\rightsquigarrow \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right).$$

Let $v = (v_1, v_2, v_3)^{\top}$ with $\|v\|_2 = 1$, and

$$c \equiv (c_1, c_2, c_3)^{\top} = \begin{pmatrix} 1 & \rho_{1n} & 0 \\ \rho_{1n} & 1 & \rho_{2n} \\ 0 & \rho_{2n} & 1 \end{pmatrix}^{-1/2} (v_1, v_2, v_3)^{\top},$$

and note that $\|c\|_2$ is bounded for all n large enough. By Cramer-Wald device, it suffices to show that

$$c^{\top} \begin{pmatrix} \frac{1}{\sqrt{\Psi}} \sum_{g \in [G]} \acute{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} \\ \frac{1}{\sqrt{\Sigma}} \left(\sum_{g \in [G]} \hat{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} + \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \right) \\ \frac{1}{\sqrt{\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \tilde{e}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \end{pmatrix} \rightsquigarrow \mathcal{N}(0, 1).$$

Denote

$$M_n = \frac{c_1}{\sqrt{\Psi}} \sum_{g \in [G]} \acute{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} + \frac{c_2}{\sqrt{\Sigma}} \left(\sum_{g \in [G]} \hat{\Pi}_{[g]}^{\top} \tilde{e}_{[g]} + \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]} \right)$$
$$+ \frac{c_3}{\sqrt{\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \tilde{e}_{[g]}^{\top} P_{[g,h]} \tilde{e}_{[h]},$$

and note that M_n can be written as sum of martingale difference array. To see this, define

$$M_{1G} = \frac{c_1}{\sqrt{\Psi}}\acute{\Pi}_{[1]}^\top \tilde{e}_{[1]} + \frac{c_2}{\sqrt{\Sigma}}\hat{\Pi}_{[1]}^\top \tilde{e}_{[1]},$$

and for $g \geqslant 2$ define

$$\begin{aligned} M_{gG} = & \frac{c_1}{\sqrt{\Psi}}\acute{\Pi}_{[g]}^\top \tilde{e}_{[g]} + \frac{c_2}{\sqrt{\Sigma}}\hat{\Pi}_{[g]}^\top \tilde{e}_{[g]} + \frac{c_2}{\sqrt{\Sigma}}\sum_{h<g}\left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]} + \tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right) \\ & + \frac{c_3}{\sqrt{\Upsilon}}\sum_{h<g}\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]} + \tilde{e}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right), \end{aligned}$$

and we have $M_n = \sum_{g\in[G]} M_{gG}$. Further define $\varepsilon_{[g]} = \left(\tilde{e}_{[g]}^\top, \tilde{V}_{[g]}^\top\right)^\top$ and the sequence of σ-fields $\mathcal{F}_{gG} = \{\varepsilon_{[1]}, \cdots, \varepsilon_{[g]}\}$ such that $\mathcal{F}_{(g-1)G} \subset \mathcal{F}_{gG}$ with $\mathcal{F}_{0G} = \{\varnothing, \Omega\}$. It is clear that $\{M_{gG}\}_{g=1}^G$ is a sequence of martingale difference array with respect to $\{\mathcal{F}_{gG}\}_{g=1}^G$. Note that

$$\sum_{g\in[G]} \mathbb{E}(M_{gG}^2) = \mathbb{E}(M_n^2) = 1$$

by the property of martingale difference array.

In order to show

$$\sum_{g\in[G]} M_{gG} \rightsquigarrow \mathcal{N}(0,1),$$

it remains to check the Lindeberg's condition

$$\sum_{g\in[G]} \mathbb{E}\left(M_{gG}^2 \mathbf{1}_{\{|M_{gG}|\geqslant c\}}\right) \to 0$$

for every $c > 0$, and the stability condition

$$\sum_{g\in[G]} \mathbb{E}\left(M_{gG}^2 | \mathcal{F}_{(g-1)G}\right) \xrightarrow{p} \sum_{g\in[G]} \mathbb{E}(M_{gG}^2).$$

Step 2: Check Lindeberg's condition. It suffices to show that

$$\sum_{g\in[G]} \mathbb{E}(M_{gG}^{(2+\delta)}) \to 0$$

for some $\delta > 0$, and we will verify that for $\delta = 2$ in what follows. We have

$$\sum_{g\in G} \mathbb{E}(M_{gG}^4) \leqslant \frac{C}{\Psi^2}\sum_{g\in[G]} \mathbb{E}\left(\acute{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)^4 + \frac{C}{\Sigma^2}\sum_{g\in[G]} \mathbb{E}\left(\hat{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)^4 + \frac{C}{\Sigma^2}\sum_{g\in[G]} \mathbb{E}\left(\sum_{h<g} \tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)^4$$
$$+ \frac{C}{\Sigma^2}\sum_{g\in[G]} \mathbb{E}\left(\sum_{h<g} \tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)^4 + \frac{C}{\Upsilon^2}\sum_{g\in[G]} \mathbb{E}\left(\sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)^4.$$

For the first term, we have

$$\begin{aligned}
&\frac{1}{\Psi^2}\sum_{g\in[G]} \mathbb{E}\left(\acute{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)^4 \\
\leqslant\ & \frac{C}{\Psi^2}\sum_{g\in[G]} \left(\acute{\Pi}_{[g]}^\top \acute{\Pi}_{[g]}\right)^2 \\
\leqslant\ & \frac{C\max_{1\leqslant g\leqslant G}\left\|\acute{\Pi}_{[g]}\right\|_2^2 \acute{\Pi}^\top\acute{\Pi}}{\left(\acute{\Pi}^\top\acute{\Pi}\right)^2} \\
=\ & o(1).
\end{aligned}$$

Here we use the fact that $\acute{\Pi}_{[g]} = z_{[g]}A_n z^\top \Pi$, so that

$$\begin{aligned}
\frac{\max_{1\leqslant g\leqslant G}\left\|\acute{\Pi}_{[g]}\right\|_2^2}{\acute{\Pi}^\top\acute{\Pi}} &= \frac{\max_{1\leqslant g\leqslant G}\left\|\acute{\Pi}_{[g]}\right\|_2^2/n}{\Pi^\top z A_n(z^\top z/n)A_n z^\top \Pi} \\
&\leqslant \frac{C\max_{1\leqslant g\leqslant G}\Pi^\top z A_n(z_{[g]}^\top z_{[g]}/n)A_n z^\top\Pi}{\|A_n z^\top \Pi\|_2^2} \\
&\leqslant \frac{C\max_{1\leqslant g\leqslant G} n_g \times \max_{i\in I_g, g\in[G]}\|z_{i,g}\|_2^2}{n} \\
&= o(1).
\end{aligned}$$

For the second term, we have

$$\begin{aligned}
&\frac{1}{\Sigma^2}\sum_{g\in[G]}\mathbb{E}\left(\hat{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)^4\\
\leqslant&\frac{C}{\Sigma^2}\sum_{g\in[G]}\left(\hat{\Pi}_{[g]}^\top \hat{\Pi}_{[g]}\right)^2\\
\leqslant&\frac{C\max_{1\leqslant g\leqslant G}\left\|\hat{\Pi}_{[g]}\right\|_2^2\hat{\Pi}^\top\hat{\Pi}}{\left(\Pi^\top\Pi+K\right)^2}\\
=&o(1)
\end{aligned}$$

as $K\to\infty$.

For the third term, we have

$$\begin{aligned}
\frac{1}{\Sigma^2}\sum_{g\in[G]}\mathbb{E}\left(\sum_{h<g}\tilde{e}_{[h]}^\top P_{[h,g]}\tilde{V}_{[g]}\right)^4 \leqslant&\frac{C}{\Sigma^2}\sum_{g\in[G]}\sum_{h<g}\mathbb{E}\left(\tilde{e}_{[h]}^\top P_{[h,g]}\tilde{V}_{[g]}\right)^4\\
&+\frac{C}{\Sigma^2}\sum_{g\in[G]}\sum_{h,k<g,h\neq k}\mathbb{E}\left(\tilde{e}_{[h]}^\top P_{[h,g]}\tilde{V}_{[g]}\right)^2\left(\tilde{e}_{[k]}^\top P_{[k,g]}\tilde{V}_{[g]}\right)^2,
\end{aligned}$$

where

$$\begin{aligned}
&\frac{1}{\Sigma^2}\sum_{g\in[G]}\sum_{h<g}\mathbb{E}\left(\tilde{e}_{[h]}^\top P_{[h,g]}\tilde{V}_{[g]}\right)^4\\
\leqslant&\frac{1}{\Sigma^2}\sum_{g\in[G]}\sum_{h<g}\lambda_{\max}^2\left(P_{[h,g]}P_{[g,h]}\right)\\
\leqslant&\frac{C}{K^2}\sum_{g\in[G]}\sum_{h<g}\operatorname{trace}\left(P_{[h,g]}P_{[g,h]}\right)\\
=&o(1)
\end{aligned}$$

and

$$\frac{1}{\Sigma^2}\sum_{g\in[G]}\sum_{h,k<g,h\neq k}\mathbb{E}\left(\tilde{e}_{[h]}^\top P_{[h,g]}\tilde{V}_{[g]}\right)^2\left(\tilde{e}_{[k]}^\top P_{[k,g]}\tilde{V}_{[g]}\right)^2$$

$$
\begin{aligned}
&\leqslant \frac{C}{\Sigma^2} \sum_{g\in[G]} \sum_{h,k<g,h\neq k} \lambda_{\max}\left(P_{[g,h]}P_{[h,g]}\right) \lambda_{\max}\left(P_{[g,k]}P_{[k,g]}\right) \\
&\leqslant \frac{C}{\Sigma^2} \sum_{g\in[G]} \sum_{h,k<g,h\neq k} \operatorname{trace}\left(P_{[g,h]}P_{[h,g]}\right) \operatorname{trace}\left(P_{[g,k]}P_{[k,g]}\right) \\
&\leqslant \frac{C}{K^2} \sum_{g\in[G]} \sum_{h<g} \operatorname{trace}\left(P_{[g,h]}P_{[h,g]}\right) \\
&= o(1).
\end{aligned}
$$

The last two terms can be handled similarly.

Step 3: Check stability condition. The variance and conditional variance can be written as

$$
\mathbb{E}\left(M_{1G}^2\right) = \mathbb{E}\left(M_{1G}^2 | \mathcal{F}_{0G}\right) = \frac{c_1^2}{\Psi} \acute{\Pi}_{[1]}^\top \Omega_1^{\tilde{e},\tilde{e}} \acute{\Pi}_{[1]} + \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \hat{\Pi}_{[1]}^\top \Omega_1^{\tilde{e},\tilde{e}} \acute{\Pi}_{[1]} + \frac{c_2^2}{\Sigma} \hat{\Pi}_{[1]}^\top \Omega_1^{\tilde{e},\tilde{e}} \hat{\Pi}_{[1]},
$$

and for $g \geqslant 2$,

$$
\begin{aligned}
\mathbb{E}\left(M_{gG}^2\right) = &\frac{c_1^2}{\Psi} \acute{\Pi}_{[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} \acute{\Pi}_{[g]} + \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \hat{\Pi}_{[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} \acute{\Pi}_{[g]} + \frac{c_2^2}{\Sigma} \hat{\Pi}_{[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} \hat{\Pi}_{[g]} \\
&+ \frac{c_2^2}{\Sigma} \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{e}} P_{[h,g]} \Omega_g^{\tilde{V},\tilde{V}} P_{[g,h]}\right) + \frac{c_2^2}{\Sigma} \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{V},\tilde{V}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]}\right) \\
&+ \frac{4c_3^2}{\Upsilon} \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{e}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]}\right) + \frac{2c_2^2}{\Sigma} \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{V}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{V}} P_{[g,h]}\right) \\
&+ \frac{4c_2c_3}{\sqrt{\Sigma\Upsilon}} \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{e}} P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} P_{[g,h]}\right) + \frac{4c_2c_3}{\sqrt{\Sigma\Upsilon}} \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{V}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}\left(M_{gG}^2 | \mathcal{F}_{(g-1)G}\right) = &\frac{c_1^2}{\Psi} \acute{\Pi}_{[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} \acute{\Pi}_{[g]} + \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \hat{\Pi}_{[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} \acute{\Pi}_{[g]} + \frac{c_2^2}{\Sigma} \hat{\Pi}_{[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} \hat{\Pi}_{[g]} \\
&+ \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \acute{\Pi}_{[g]} + \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{V}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} \acute{\Pi}_{[g]} \\
&+ \frac{4c_1c_3}{\sqrt{\Psi\Upsilon}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \acute{\Pi}_{[g]} + \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \hat{\Pi}_{[g]}
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{V}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} \hat{\Pi}_{[g]} + \frac{4c_1c_3}{\sqrt{\Psi\Upsilon}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \hat{\Pi}_{[g]} \\
&+ \frac{c_2^2}{\Sigma} \sum_{h,k<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{V}} P_{[g,k]} \tilde{e}_{[k]} + \frac{c_2^2}{\Sigma} \sum_{h,k<g} \tilde{V}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]} \tilde{V}_{[k]} \\
&+ \frac{4c_3^2}{\Upsilon} \sum_{h,k<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]} \tilde{e}_{[k]} + \frac{2c_2^2}{\Sigma} \sum_{h,k<g} e_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} P_{[g,k]} \tilde{V}_{[k]} \\
&+ \frac{4c_2c_3}{\sqrt{\Sigma\Upsilon}} \sum_{h,k<g} V_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]} \tilde{e}_{[k]} + \frac{4c_2c_3}{\sqrt{\Sigma\Upsilon}} \sum_{h,k<g} e_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} P_{[g,k]} \tilde{e}_{[k]}.
\end{aligned}
$$

We thus obtain

$$
\begin{aligned}
&\mathbb{E}\left(M_{gG}^2 | \mathcal{F}_{(g-1)G}\right) - \mathbb{E}\left(M_{gG}^2\right) \\
&= \left\{ \begin{array}{c}
\frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \acute{\Pi}_{[g]} + \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{V}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} \acute{\Pi}_{[g]} + \frac{4c_1c_3}{\sqrt{\Psi\Upsilon}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \acute{\Pi}_{[g]} \\
+ \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \hat{\Pi}_{[g]} + \frac{2c_1c_2}{\sqrt{\Psi\Sigma}} \sum_{h<g} \tilde{V}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} \hat{\Pi}_{[g]} + \frac{4c_1c_3}{\sqrt{\Psi\Upsilon}} \sum_{h<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} \hat{\Pi}_{[g]}
\end{array} \right\} \\
&+ \left\{ \begin{array}{c}
\frac{c_2^2}{\Sigma} \left[\sum_{h,k<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{V}} P_{[g,k]} \tilde{e}_{[k]} - \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{e}} P_{[h,g]} \Omega_g^{\tilde{V},\tilde{V}} P_{[g,h]} \right) \right] \\
+ \frac{c_2^2}{\Sigma} \left[\sum_{h,k<g} \tilde{V}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]} \tilde{V}_{[k]} - \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{V},\tilde{V}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]} \right) \right] \\
+ \frac{4c_3^2}{\Upsilon} \left[\sum_{h,k<g} \tilde{e}_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]} \tilde{e}_{[k]} - \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{e}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]} \right) \right] \\
+ \frac{2c_2^2}{\Sigma} \left[\sum_{h,k<g} e_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} P_{[g,k]} \tilde{V}_{[k]} - \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{V}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{V}} P_{[g,h]} \right) \right] \\
+ \frac{4c_2c_3}{\sqrt{\Sigma\Upsilon}} \left[\sum_{h,k<g} V_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]} \tilde{e}_{[k]} - \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{e}} P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} P_{[g,h]} \right) \right] \\
+ \frac{4c_2c_3}{\sqrt{\Sigma\Upsilon}} \left[\sum_{h,k<g} e_{[h]}^\top P_{[h,g]} \Omega_g^{\tilde{V},\tilde{e}} P_{[g,k]} \tilde{e}_{[k]} - \sum_{h<g} \operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{V}} P_{[h,g]} \Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]} \right) \right]
\end{array} \right\} \\
&\equiv M_{gG}^{(1)} + M_{gG}^{(2)},
\end{aligned}
$$

and it suffices to show that

$$
\begin{aligned}
M_n^{(1)} &\equiv \sum_{g \in [G]} M_{gG}^{(1)} = o_P(1), \\
M_n^{(2)} &\equiv \sum_{g \in [G]} M_{gG}^{(2)} = o_P(1).
\end{aligned}
$$

Consider first $M_n^{(1)}$, we shall only compute the sum of the first term in $M_{gG}^{(1)}$, as the other terms can be handled similarly. Recall that $\tilde{P}$ is the block lower triangular matrix corresponding to $P-\bar{P}$, we have

$$\mathbb{E}\left(\frac{2c_1c_2}{\sqrt{\Psi\Sigma}}\sum_{g\in[G]}\sum_{h<g}\tilde{e}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{V},\tilde{e}}\acute{\Pi}_{[g]}\right)=0,$$

and

$$\begin{aligned}\mathbb{V}\left(\frac{2c_1c_2}{\sqrt{\Psi\Sigma}}\sum_{g\in[G]}\sum_{h<g}\tilde{e}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{V},\tilde{e}}\acute{\Pi}_{[g]}\right)&=\frac{4c_1^2c_2^2}{\Psi\Sigma}\mathbb{E}\left(\tilde{e}^{\top}\tilde{P}^{\top}\Omega_{\tilde{V},\tilde{e}}\acute{\Pi}\right)^2\\&\leqslant\frac{C\lambda_{\max}\left(\tilde{P}\tilde{P}^{\top}\right)\lambda_{\max}\left(\Omega_{\tilde{e},\tilde{V}}\Omega_{\tilde{V},\tilde{e}}\right)\acute{\Pi}^{\top}\acute{\Pi}}{(\acute{\Pi}^{\top}\acute{\Pi})K}\\&\leqslant\frac{C\lambda_{\max}\left(\tilde{P}\tilde{P}^{\top}\right)}{K}\\&\leqslant\frac{C\left\|\tilde{P}\tilde{P}^{\top}\right\|_F}{K}\\&=o(1),\end{aligned}$$

where we use Lemma B.2 in the last equality.

Now consider $M_n^{(2)}$, we shall only compute the sum of the first term in $M_{gG}^{(2)}$, as the other terms can be handled similarly. We have

$$\mathbb{E}\left(\frac{c_2^2}{\Sigma}\sum_{g\in[G]}\left[\sum_{h,k<g}\tilde{e}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{V},\tilde{V}}P_{[g,k]}\tilde{e}_{[k]}-\sum_{h<g}\operatorname{trace}\left(\Omega_h^{\tilde{e},\tilde{e}}P_{[h,g]}\Omega_g^{\tilde{V},\tilde{V}}P_{[g,h]}\right)\right]\right)=0,$$

and

$$\begin{aligned}&\mathbb{V}\left(\frac{c_2^2}{\Sigma}\sum_{g\in[G]}\left[\sum_{h,k<g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\tilde{V}_{[k]}-\sum_{h<g}\operatorname{trace}\left(\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\right]\right)\\&\leqslant\frac{C}{\Sigma^2}\mathbb{E}\left(\sum_{g\in[G]}\sum_{h<g}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\tilde{V}_{[h]}-\operatorname{trace}\left(\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\right)\right)^2\end{aligned}$$

$$+\frac{C}{\Sigma^2}\mathbb{E}\left(\sum_{g\in[G]}\sum_{h,k<g,h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\tilde{V}_{[k]}\right)^2.$$

For the first term, we have

$$\begin{aligned}
&\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g\in[G]}\sum_{h<g}\left(\tilde{V}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\tilde{V}_{[h]}-\text{trace}\left(\Omega_h^{\tilde{V},\tilde{V}}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\right)\right)^2\\
&=\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{h\in[G]}\tilde{V}_{[h]}^{\top}\left(\sum_{g>h}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\tilde{V}_{[h]}-\text{trace}\left(\Omega_h^{\tilde{V},\tilde{V}}\left(\sum_{g>h}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\right)\right)^2\\
&\leqslant\frac{1}{\Sigma^2}\sum_{h\in[G]}\mathbb{E}\left(\tilde{V}_{[h]}^{\top}\left(\sum_{g>h}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,h]}\right)\tilde{V}_{[h]}\right)^2\\
&\leqslant\frac{1}{\Sigma^2}\sum_{h\in[G]}\lambda_{\max}^2\left(\left(\tilde{P}^{\top}\Omega_{\tilde{e}}\tilde{P}\right)_{[h,h]}\right)\\
&\leqslant\frac{C}{\Sigma^2}\sum_{h\in[G]}\text{trace}\left(\left(\tilde{P}^{\top}\Omega_{\tilde{e}}\tilde{P}\right)_{[h,h]}\right)\\
&\leqslant\frac{C}{\Sigma^2}\text{trace}\left(\tilde{P}^{\top}\Omega_{\tilde{e}}\tilde{P}\right)\\
&\leqslant\frac{C}{K^2}\text{trace}\left(\tilde{P}^{\top}\tilde{P}\right)\\
&=o(1),
\end{aligned}$$

where we use the fact that

$$\text{trace}\left(\tilde{P}^{\top}\tilde{P}\right)=||\tilde{P}||_F^2\leqslant||P||_F^2=O(K).$$

For the second term, we have

$$\begin{aligned}
&\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{g\in[G]}\sum_{h,k<g,h\neq k}\tilde{V}_{[h]}^{\top}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\tilde{V}_{[k]}\right)^2\\
&=\frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{h,k\in[G]^2,h\neq k}\tilde{V}_{[h]}^{\top}\left(\sum_{g>h\vee k}P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}}P_{[g,k]}\right)\tilde{V}_{[k]}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\Sigma^2}\mathbb{E}\left(\sum_{h,k\in[G]^2,h<k} \tilde{V}_{[h]}^\top \left(\sum_{g>k} P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]}\right)\tilde{V}_{[k]}\right. \\
&\qquad \left. + \sum_{h,k\in[G]^2,h>k} \tilde{V}_{[h]}^\top \left(\sum_{g>h} P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]}\right)\tilde{V}_{[k]}\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h,k\in[G]^2,h\neq k} \mathbb{E}\left(\tilde{V}_{[h]}^\top \left(\sum_{g>h\vee k} P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]}\right)\tilde{V}_{[k]}\right)^2 \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h,k\in[G]^2,h\neq k} \operatorname{trace}\left(\left(\sum_{g>h\vee k} P_{[h,g]}\Omega_g^{\tilde{e},\tilde{e}} P_{[g,k]}\right)\left(\sum_{g>h\vee k} P_{[k,g]}\Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]}\right)\right) \\
&\leqslant \frac{C}{\Sigma^2}\sum_{h,k\in[G]^2} \operatorname{trace}\left(\left(\tilde{P}^\top\Omega_{\tilde{e}}\tilde{P}\right)_{[h,k]}\left(\tilde{P}^\top\Omega_{\tilde{e}}\tilde{P}\right)_{[k,h]}\right) \\
&\leqslant \frac{C}{K^2}\operatorname{trace}\left(\tilde{P}^\top\Omega_{\tilde{e}}\tilde{P}\tilde{P}^\top\Omega_{\tilde{e}}\tilde{P}\right) \\
&\leqslant \frac{C\left\|\tilde{P}\tilde{P}^\top\right\|_F^2}{K^2} \\
&= o(1),
\end{aligned}
$$

where we use Lemma B.2 in the last equality. This concludes the proof. □

C.11 Proof of Lemma B.11

For the first result, for $j \in [d_z]$, let v_j be the d_z-dimensional unit vector with j-th element one and other elements zero, and denote $\bar{z}_j = z\Omega^{-1/2}v_j$, then

$$
\begin{aligned}
\bar{z}_j^\top \bar{z}_j &= v_j^\top(\Omega/n)^{-1/2}(z^\top z/n)(\Omega/n)^{-1/2}v_j = O(1), \\
\max_{g\in[G]} \bar{z}_{j,[g]}^\top \bar{z}_{j,[g]} &= \max_{g\in[G]} v_j^\top(\Omega/n)^{-1/2}(z_{[g]}^\top z_{[g]}/n)(\Omega/n)^{-1/2}v_j = o(1).
\end{aligned}
$$

It follows that

$$
\left|\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\mathbb{E}\left[\left(\bar{z}_{j,[g]}^\top\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^\top\tilde{e}_{[g]}\right)\right]\right| \leqslant \left(\sum_{g\in[G]}\mathbb{E}\left(\bar{z}_{j,[g]}^\top\tilde{e}_{[g]}\right)^2\right)^{1/2}\left(\frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\hat{\Pi}_{[g]}^\top\tilde{e}_{[g]}\right)^2\right)^{1/2} = O(1),
$$

and as d_z is fixed, we obtain

$$\frac{1}{\sqrt{\Sigma}}\Omega^{-1/2}\sum_{g\in[G]}\mathbb{E}\left[\left(z_{[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right]=O(1). \tag{C.32}$$

In addition, let $\hat{V}=M_W(P-\bar{P})V=Q\tilde{V}$, we have

$$\begin{aligned}\frac{1}{\sqrt{\Sigma}}\Omega^{-1/2}\sum_{g\in[G]}\left[\left(z_{[g]}^{\top}\hat{e}_{[g]}\right)\left(\hat{X}_{[g]}^{\top}\hat{e}_{[g]}\right)\right]&=\frac{1}{\sqrt{\Sigma}}\Omega^{-1/2}\sum_{g\in[G]}\left[\left(z_{[g]}^{\top}\hat{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\hat{e}_{[g]}\right)\right]\\&+\frac{1}{\sqrt{\Sigma}}\Omega^{-1/2}\sum_{g\in[G]}\left[\left(z_{[g]}^{\top}\hat{e}_{[g]}\right)\left(\hat{V}_{[g]}^{\top}\hat{e}_{[g]}\right)\right].\end{aligned}$$

Thus, it suffices to show that, for any j,

$$\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left[\left(\bar{z}_{j,[g]}^{\top}\hat{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\hat{e}_{[g]}\right)\right]-\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\mathbb{E}\left[\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right]=o_P(1), \tag{C.33}$$

$$\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left[\left(\bar{z}_{j,[g]}^{\top}\hat{e}_{[g]}\right)\left(\hat{V}_{[g]}^{\top}\hat{e}_{[g]}\right)\right]=o_P(1). \tag{C.34}$$

For (C.33), the left-hand side can be written as

$$\begin{aligned}&\frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\hat{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\hat{e}_{[g]}\right)-\sum_{g\in[G]}\mathbb{E}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right)\\&=\underbrace{\frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)-\sum_{g\in[G]}\mathbb{E}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right)}_{R_{27}}\\&+\underbrace{\frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}e_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}e_{[g]}\right)-\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right)}_{R_{28}}\\&+\underbrace{\frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\hat{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\hat{e}_{[g]}\right)-\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}e_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}e_{[g]}\right)\right)}_{R_{29}}.\end{aligned}$$

For R_{27}, it has mean zero and

$$\begin{aligned}
\mathbb{V}\left(R_{27}\right) &= \frac{1}{\Sigma}\mathbb{E}\left(\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)-\sum_{g\in[G]}\mathbb{E}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right)^{2} \\
&\leqslant \frac{1}{\Sigma}\sum_{g\in[G]}\mathbb{E}\left(\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)\right)^{2} \\
&\leqslant \frac{C\max_{1\leqslant g\leqslant G}\left\|\hat{\Pi}_{[g]}\right\|_{2}^{2}\bar{z}_{j}^{\top}\bar{z}_{j}}{\left(\Pi^{\top}\Pi+K\right)} \\
&= o(1).
\end{aligned}$$

For R_{28}, we have

$$\begin{aligned}
R_{28} = &\underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)\left(\hat{\Pi}_{[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)}_{R_{28,1}} \\
&-\underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)}_{R_{28,2}} \\
&-\underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)}_{R_{28,3}},
\end{aligned}$$

where

$$\begin{aligned}
\left|R_{28,1}\right| &\leqslant \max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_{2}^{2}\times\sqrt{\bar{z}_{j}^{\top}\bar{z}_{j}}\times\sqrt{\frac{\hat{\Pi}^{\top}\hat{\Pi}}{\Sigma}}=o_{P}(1), \\
\left|R_{28,2}\right| &\leqslant \max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_{2}\times\sqrt{\bar{z}_{j}^{\top}\bar{z}_{j}}\times\sqrt{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\hat{\Pi}_{[g]}^{\top}\tilde{e}_{[g]}\right)^{2}}=o_{P}(1), \\
\left|R_{28,3}\right| &\leqslant \max_{1\leqslant g\leqslant G}\left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_{2}\times\sqrt{\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)^{2}}\times\sqrt{\frac{\hat{\Pi}^{\top}\hat{\Pi}}{\Sigma}}=o_{P}(1),
\end{aligned}$$

whence $R_{28} = o_P(1)$. For R_{29}, we have

$$
\begin{aligned}
R_{29} = \hat{\Delta}^2 \times & \underbrace{\frac{1}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top X_{[g]}\right) \left(\hat{\Pi}_{[g]}^\top X_{[g]}\right)}_{R_{29,1}} \\
- \hat{\Delta} \times & \underbrace{\frac{1}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top X_{[g]}\right) \left(\hat{\Pi}_{[g]}^\top e_{[g]}\right)}_{R_{29,2}} \\
- \hat{\Delta} \times & \underbrace{\frac{1}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top e_{[g]}\right) \left(\hat{\Pi}_{[g]}^\top e_{[g]}\right)}_{R_{29,3}},
\end{aligned}
$$

where $\hat{\Delta} = o_P(1)$ and

$$
\begin{aligned}
|R_{29,1}| &\leqslant \sqrt{\sum_{g \in [G]} \bar{z}_{j,[g]}^\top \bar{z}_{j,[g]} X_{[g]}^\top X_{[g]}} \times \sqrt{\frac{1}{\Sigma} \sum_{g \in [G]} \hat{\Pi}_{[g]}^\top \hat{\Pi}_{[g]} X_{[g]}^\top X_{[g]}} = O_P(1), \\
|R_{29,2}| &\leqslant \sqrt{\sum_{g \in [G]} \bar{z}_{j,[g]}^\top \bar{z}_{j,[g]} X_{[g]}^\top X_{[g]}} \times \sqrt{\frac{1}{\Sigma} \sum_{g \in [G]} \hat{\Pi}_{[g]}^\top \hat{\Pi}_{[g]} e_{[g]}^\top e_{[g]}} = O_P(1), \\
|R_{29,3}| &\leqslant \sqrt{\sum_{g \in [G]} \bar{z}_{j,[g]}^\top \bar{z}_{j,[g]} e_{[g]}^\top e_{[g]}} \times \sqrt{\frac{1}{\Sigma} \sum_{g \in [G]} \hat{\Pi}_{[g]}^\top \hat{\Pi}_{[g]} e_{[g]}^\top e_{[g]}} = O_P(1),
\end{aligned}
$$

since $\max_{g \in [G]} \mathbb{E}\left(X_{[g]}^\top X_{[g]}\right) = O(1)$ and $\max_{g \in [G]} \mathbb{E}\left(e_{[g]}^\top e_{[g]}\right) = O(1)$ by Lemma B.1. It follows that $R_{29} = o_P(1)$.

For (C.34), the left-hand side can be written as

$$
\begin{aligned}
&\frac{1}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top \hat{e}_{[g]}\right) \left(\hat{V}_{[g]}^\top \hat{e}_{[g]}\right) \\
&= \frac{\hat{\Delta}^2}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top X_{[g]}\right) \left(\hat{V}_{[g]}^\top X_{[g]}\right) - \frac{\hat{\Delta}}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top X_{[g]}\right) \left(\hat{V}_{[g]}^\top e_{[g]}\right) \\
&- \frac{\hat{\Delta}}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top e_{[g]}\right) \left(\hat{V}_{[g]}^\top X_{[g]}\right) + \frac{1}{\sqrt{\Sigma}} \sum_{g \in [G]} \left(\bar{z}_{j,[g]}^\top e_{[g]}\right) \left(\hat{V}_{[g]}^\top e_{[g]}\right).
\end{aligned}
$$

Here we only show that the last term is $o_P(1)$, which is most difficult since it does not involve

$\hat{\Delta}$. We have

$$\underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}e_{[g]}\right)\left(\hat{V}_{[g]}^{\top}e_{[g]}\right)}_{R_{30}} = \underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{V}_{[g]}^{\top}\tilde{e}_{[g]}\right)}_{R_{30,1}}$$

$$- \underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)\left(\hat{V}_{[g]}^{\top}\tilde{e}_{[g]}\right)}_{R_{30,2}}$$

$$- \underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}\tilde{e}_{[g]}\right)\left(\hat{V}_{[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)}_{R_{30,3}}$$

$$+ \underbrace{\frac{1}{\sqrt{\Sigma}}\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)\left(\hat{V}_{[g]}^{\top}W_{[g]}\hat{\gamma}_{\tilde{e}}\right)}_{R_{30,4}}.$$

Note that

$$\hat{V}_{[g]} = \sum_{h\in[G],h\neq g}P_{[g,h]}\tilde{V}_{[h]} + \sum_{h\in[G]}P_{W,[g,h]}P_{[h,h]}\tilde{V}_{[h]}$$
$$+ \sum_{h\in[G]}P_{[g,g]}P_{W,[g,h]}\tilde{V}_{[h]} - \sum_{h\in[G]}\left(\sum_{k\in[G]}P_{W,[g,k]}P_{[k,k]}P_{W,[k,h]}\right)\tilde{V}_{[h]}.$$

Therefore, we have

$$\frac{1}{\Sigma}\sum_{g\in[G]}\left(\hat{V}_{[g]}^{\top}\tilde{e}_{[g]}\right)^2 \leqslant \frac{C}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G],h\neq g}\tilde{V}_{[h]}^{\top}P_{[h,g]}\tilde{e}_{[g]}\right)^2$$
$$+ \frac{C}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G]}\tilde{V}_{[h]}^{\top}P_{[h,h]}P_{W,[h,g]}\tilde{e}_{[g]}\right)^2$$
$$+ \frac{C}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G]}\tilde{V}_{[h]}^{\top}P_{W,[h,g]}P_{[g,g]}\tilde{e}_{[g]}\right)^2$$
$$+ \frac{C}{\Sigma}\sum_{g\in[G]}\left(\sum_{h\in[G]}\tilde{V}_{[h]}^{\top}\left(\sum_{k\in[G]}P_{W,[h,k]}P_{[k,k]}P_{W,[k,g]}\right)\tilde{e}_{[g]}\right)^2$$

$$
\begin{aligned}
&= \frac{C}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top P_{[h,g]} \tilde{e}_{[g]} \right)^2 + o_P(1) \\
&= \frac{C}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G], h \neq g} \tilde{V}_{[h]}^\top P_{[h,g]} \tilde{e}_{[g]} \right)^2 + o_P(1) \\
&= O_P(1),
\end{aligned}
$$

by Lemma B.4 and the facts that

$$
\begin{aligned}
\frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G]} \tilde{V}_{[h]}^\top P_{[h,h]} P_{W,[h,g]} \tilde{e}_{[g]} \right)^2 &= \frac{1}{\Sigma} \sum_{g,h \in [G]^2} \mathbb{E} \left(\tilde{V}_{[h]}^\top P_{[h,h]} P_{W,[h,g]} \tilde{e}_{[g]} \right)^2 \\
&\leqslant \frac{C \text{trace} \left(P_W \bar{P}^2 P_W \right)}{\Sigma} \\
&= o(1),
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{\Sigma} \sum_{g \in [G]} \mathbb{E} \left(\sum_{h \in [G]} \tilde{V}_{[h]}^\top P_{W,[h,g]} P_{[g,g]} \tilde{e}_{[g]} \right)^2 &= \frac{1}{\Sigma} \sum_{g,h \in [G]^2} \mathbb{E} \left(\tilde{V}_{[h]}^\top P_{W,[h,g]} P_{[g,g]} \tilde{e}_{[g]} \right)^2 \\
&\leqslant \frac{C \text{trace} \left(\bar{P} P_W^2 \bar{P} \right)}{\Sigma} \\
&= o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{\Sigma} \sum_{g \in [G]} \left(\sum_{h \in [G]} \tilde{V}_{[h]}^\top \left(\sum_{k \in [G]} P_{W,[h,k]} P_{[k,k]} P_{W,[k,g]} \right) \tilde{e}_{[g]} \right)^2 \\
&= \frac{1}{\Sigma} \sum_{g,h \in [G]^2} \left(\tilde{V}_{[h]}^\top \left(\sum_{k \in [G]} P_{W,[h,k]} P_{[k,k]} P_{W,[k,g]} \right) \tilde{e}_{[g]} \right)^2 \\
&\leqslant \frac{C \text{trace} \left(P_W \bar{P} P_W^2 \bar{P} P_W \right)}{\Sigma} \\
&= o(1).
\end{aligned}
$$

This implies

$$|R_{30,2}| \leqslant \max_{1\leqslant g\leqslant G} \left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2 \times \sqrt{\bar{z}_j^\top \bar{z}_j} \times \sqrt{\frac{1}{\Sigma}\sum_{g\in[G]}\left(\hat{V}_{[g]}^\top \tilde{e}_{[g]}\right)^2} = o_P(1).$$

In addition, we have

$$|R_{30,3}| \leqslant \max_{1\leqslant g\leqslant G} \left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2 \times \sqrt{\sum_{g\in[G]}\left(\bar{z}_{j,[g]}^\top \tilde{e}_{[g]}\right)^2} \times \sqrt{\frac{\tilde{V}^\top Q^2 \tilde{V}}{\Sigma}} = o_P(1),$$

$$|R_{30,4}| \leqslant \max_{1\leqslant g\leqslant G} \left\|W_{[g]}\hat{\gamma}_{\tilde{e}}\right\|_2^2 \times \sqrt{\bar{z}_j^\top \bar{z}_j} \times \sqrt{\frac{\tilde{V}^\top Q^2 \tilde{V}}{\Sigma}} = o_P(1).$$

It follows that

$$R_{30} = R_{30,1} + o_P(1),$$

and for $R_{30,1}$, note that similar to the proof above, we also have

$$R_{30,1} = \frac{1}{\sqrt{\Sigma}} \sum_{g,h\in[G]^2 g\neq h} \left(\bar{z}_{j,[g]}^\top \tilde{e}_{[g]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right) + o_P(1),$$

where

$$\mathbb{E}\left(\frac{1}{\sqrt{\Sigma}} \sum_{g,h\in[G]^2 g\neq h} \left(z_{[g]}^\top \tilde{e}_{[g]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)\right) = 0,$$

and

$$\begin{aligned}&\mathbb{V}\left(\frac{1}{\sqrt{\Sigma}} \sum_{g,h\in[G]^2 g\neq h} \left(\bar{z}_{j,[g]}^\top \tilde{e}_{[g]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)\right)\\ &= \frac{1}{\Sigma}\mathbb{E}\left(\sum_{g,h\in[G]^2 g\neq h} \left(\bar{z}_{j,[g]}^\top \tilde{e}_{[g]}\right)\left(\tilde{V}_{[h]}^\top P_{[h,g]}\tilde{e}_{[g]}\right)\right)^2\end{aligned}$$

$$
\begin{aligned}
&\leqslant \frac{C}{\Sigma}\mathbb{E}\left(\sum_{g,h\in[G]^2 g\neq h} \bar{z}_{j,[g]}^\top \left(\tilde{e}_{[g]}\tilde{e}_{[g]}^\top - \Omega_g^{\tilde{e},\tilde{e}}\right) P_{[g,h]}\tilde{V}_{[h]}\right)^2 \\
&+ \frac{C}{\Sigma}\mathbb{E}\left(\sum_{g,h\in[G]^2 g\neq h} \bar{z}_{j,[g]}^\top \Omega_g^{\tilde{e},\tilde{e}} P_{[g,h]}\tilde{V}_{[h]}\right)^2 \\
&\leqslant \frac{C}{\Sigma}\sum_{g,h\in[G]^2 g\neq h} \mathbb{E}\left(\bar{z}_{j,[g]}^\top \tilde{e}_{[g]}\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{V}_{[h]}\right)^2 \\
&+ \frac{C}{\Sigma}\mathbb{E}\left(\bar{z}_j^\top \Omega_{\tilde{e}}(P-\bar{P})\tilde{V}\right)^2 \\
&\leqslant \frac{C\max_{g\in[G]}\left\|\bar{z}_{j,[g]}\right\|_2^2}{\Sigma}\sum_{g,h\in[G]^2 g\neq h} \operatorname{trace}\left(P_{[g,h]}P_{[h,g]}\right) + \frac{C\bar{z}_j^\top\bar{z}_j}{\Sigma} \\
&= o(1).
\end{aligned}
$$

For the second result, we have

$$
\begin{aligned}
&\frac{1}{\sqrt{\Sigma\Upsilon}}\left[\sum_{g,h\in[G]^2, g\neq h}\left(X_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right)\left(\hat{e}_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right)\right. \\
&\qquad\left. - \sum_{g,h\in[G]^2, g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\right] \\
&= \frac{1}{\sqrt{\Sigma\Upsilon}}\sum_{g,h\in[G]^2, g\neq h}\left(\Pi_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right)\left(\hat{e}_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right) \\
&+ \frac{1}{\sqrt{\Sigma\Upsilon}}\left[\sum_{g,h\in[G]^2, g\neq h}\left(V_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right)\left(\hat{e}_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right)\right. \\
&\qquad\left. - \sum_{g,h\in[G]^2, g\neq h}\mathbb{E}\left(\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\left(\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)\right].
\end{aligned}
\tag{C.35}
$$

The first term on the RHS of (C.35) can be written as

$$
\frac{1}{\sqrt{\Sigma\Upsilon}}\sum_{g,h\in[G]^2, g\neq h}\left(\Pi_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right)\left(\hat{e}_{[g]}^\top P_{[g,h]}\hat{e}_{[h]}\right)
$$

$$= \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \left(\tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right)}_{R_{31}}$$

$$+ \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \left(\sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} e_{[h]}\right) \left(e_{[g]}^\top P_{[g,h]} e_{[h]}\right) - \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \left(\tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \right)}_{R_{32}}$$

$$+ \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \left(\sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) \left(\hat{e}_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) - \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} e_{[h]}\right) \left(e_{[g]}^\top P_{[g,h]} e_{[h]}\right) \right)}_{R_{33}}.$$

By using the same argument as in the proof of Lemma B.6, we can show that $R_{31} = o_P(1)$ and $R_{32} = o_P(1)$. For R_{33}, we have

$$R_{33} = -\hat{\Delta}^3 \times \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} X_{[h]}\right) \left(X_{[g]}^\top P_{[g,h]} X_{[h]}\right)}_{R_{33,1}}$$

$$+ \hat{\Delta}^2 \times \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} e_{[h]}\right) \left(X_{[g]}^\top P_{[g,h]} X_{[h]}\right)}_{R_{33,2}}$$

$$+ \hat{\Delta}^2 \times \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} X_{[h]}\right) \left(e_{[g]}^\top P_{[g,h]} X_{[h]}\right)}_{R_{33,3}}$$

$$+ \hat{\Delta}^2 \times \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} X_{[h]}\right) \left(X_{[g]}^\top P_{[g,h]} e_{[h]}\right)}_{R_{33,4}}$$

$$- \hat{\Delta} \times \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} X_{[h]}\right) \left(e_{[g]}^\top P_{[g,h]} e_{[h]}\right)}_{R_{33,5}}$$

$$- \hat{\Delta} \times \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(\Pi_{[g]}^\top P_{[g,h]} e_{[h]}\right) \left(X_{[g]}^\top P_{[g,h]} e_{[h]}\right)}_{R_{33,6}}$$

$$-\hat{\Delta} \times \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \left(\Pi_{[g]}^\top P_{[g,h]} e_{[h]}\right) \left(e_{[g]}^\top P_{[g,h]} X_{[h]}\right)}_{R_{33,7}},$$

and by using a similar argument as in the proof of Lemma B.6, we can show that

$$R_{33,i} = O_P(1), \quad i = 1, \ldots, 7,$$

which implies that $R_{33} = o_P(1)$.

The second term on the RHS of (C.35) can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{\Sigma\Upsilon}} \left(\sum_{g,h \in [G]^2, g \neq h} \left(V_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) \left(\hat{e}_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) - \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \left(\tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \right) \\
&= \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \left(\sum_{g,h \in [G]^2, g \neq h} \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \left(\tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) - \sum_{g,h \in [G]^2, g \neq h} \mathbb{E} \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \left(\tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \right)}_{R_{34}} \\
&+ \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \left(\sum_{g,h \in [G]^2, g \neq h} \left(V_{[g]}^\top P_{[g,h]} e_{[h]}\right) \left(e_{[g]}^\top P_{[g,h]} e_{[h]}\right) - \sum_{g,h \in [G]^2, g \neq h} \left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \left(\tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \right)}_{R_{35}} \\
&+ \underbrace{\frac{1}{\sqrt{\Sigma\Upsilon}} \left(\sum_{g,h \in [G]^2, g \neq h} \left(V_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) \left(\hat{e}_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) - \sum_{g,h \in [G]^2, g \neq h} \left(V_{[g]}^\top P_{[g,h]} e_{[h]}\right) \left(e_{[g]}^\top P_{[g,h]} e_{[h]}\right) \right)}_{R_{36}}.
\end{aligned}$$

By using the same argument as in the proof of Lemma B.6, we can show that $R_{34} = o_P(1)$ and $R_{35} = o_P(1)$. In addition, we can show that $R_{36} = o_P(1)$ as in the proof for R_{33}. This concludes the proof. □

C.12 Proof of Lemma B.12

If Assumptions 1–4 hold, then by Lemma B.8, we have $\hat{\beta} \xrightarrow{p} \beta$. Consider first $\hat{\rho}_1$. It suffices to show that $\hat{\rho}_1 - \rho_{1n} = o_P(1)$, and we note that $\hat{\rho}_1$ can be written as

$$
\begin{aligned}
\hat{\rho}_1 &= \frac{1}{\sqrt{\hat{\Psi}\hat{\Sigma}}} \sum_{g \in [G]} \left[\left(\acute{X}_{[g]}^{\top} \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^{\top} \hat{e}_{[g]} \right) \right] \\
&= \frac{1}{\sqrt{\hat{\Psi}\hat{\Sigma}}} \sum_{g \in [G]} \left[\left((z\hat{A}_n z^{\top} X)_{[g]}^{\top} \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^{\top} \hat{e}_{[g]} \right) \right] \\
&= \frac{1}{\sqrt{\hat{\Psi}\hat{\Sigma}}} X^{\top} z \hat{A}_n \sum_{g \in [G]} \left[\left(z_{[g]}^{\top} \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^{\top} \hat{e}_{[g]} \right) \right] \\
&= \sqrt{\frac{\Sigma}{\hat{\Sigma}}} \times \frac{1}{\sqrt{X^{\top} z \hat{A}_n \hat{\Omega} \hat{A}_n z^{\top} X}} X^{\top} z \hat{A}_n \hat{\Omega}^{1/2} \\
&\quad \times \hat{\Omega}^{-1/2} \Omega^{1/2} \times \frac{1}{\sqrt{\Sigma}} \Omega^{-1/2} \sum_{g \in [G]} \left[\left(z_{[g]}^{\top} \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^{\top} \hat{e}_{[g]} \right) \right].
\end{aligned}
$$

By Lemma B.7, we have

$$
\begin{aligned}
\sqrt{\frac{\Sigma}{\hat{\Sigma}}} &= 1 + o_P(1), \\
\hat{\Omega}^{-1/2} \Omega^{1/2} &= I_{d_z} + o_P(1),
\end{aligned}
$$

and by (C.17) and (C.18) we have

$$
\begin{aligned}
&\frac{1}{\sqrt{X^{\top} z \hat{A}_n \hat{\Omega} \hat{A}_n z^{\top} X}} X^{\top} z \hat{A}_n \hat{\Omega}^{1/2} \\
&= \frac{1}{\sqrt{(X^{\top} z / r_n)(\hat{A}_n / \lambda_n)(\hat{\Omega}/n)(\hat{A}_n/\lambda_n)(z^{\top} X / r_n)}} (X^{\top} z / r_n)(\hat{A}_n / \lambda_n)(\hat{\Omega}/n)^{1/2} \\
&= \frac{1}{\sqrt{(\Pi^{\top} z / r_n)(A_n / \lambda_n)(\Omega/n)(A_n/\lambda_n)(z^{\top} \Pi / r_n)}} (\Pi^{\top} z / r_n)(A_n / \lambda_n)(\Omega/n)^{1/2} + o_P(1) \\
&= \frac{1}{\sqrt{\Pi^{\top} z A_n \Omega A_n z^{\top} \Pi}} \Pi^{\top} z A_n \Omega^{1/2} + o_P(1).
\end{aligned}
$$

Further recall that

$$\rho_{1n} = \frac{1}{\sqrt{\Psi\Sigma}} \sum_{g\in[G]} \mathbb{E}\left[\left(\acute{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)\right]$$
$$= \frac{1}{\sqrt{\Pi^\top z A_n \Omega A_n z^\top \Pi}} \Pi^\top z A_n \Omega^{1/2} \times \frac{1}{\sqrt{\Sigma}} \Omega^{-1/2} \sum_{g\in[G]} \mathbb{E}\left[\left(z_{[g]}^\top \tilde{e}_{[g]}\right)\left(\hat{\Pi}_{[g]}^\top \tilde{e}_{[g]}\right)\right].$$

The consistency of $\hat{\rho}_1$ then follows by (C.32) and Lemma B.11.

Next, consider $\hat{\rho}_2$. It suffices to show that $\hat{\rho}_2 - \rho_{2n} = o_P(1)$, we have

$$\hat{\rho}_2 - \rho_{2n} = \frac{2}{\sqrt{\Sigma\Upsilon}} \left[\sum_{g,h\in[G]^2, g\neq h} \left(X_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right)\left(\hat{e}_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) \right.$$
$$\left. - \sum_{g,h\in[G]^2, g\neq h} \mathbb{E}\left(\tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right)\left(\tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]}\right) \right]$$
$$+ \left(\sqrt{\frac{\Sigma\Upsilon}{\hat{\Sigma}\hat{\Upsilon}}} - 1\right) \times \left(\frac{2}{\sqrt{\Sigma\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \left(X_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right)\left(\hat{e}_{[g]}^\top P_{[g,h]} \hat{e}_{[h]}\right) \right).$$

The consistency of $\hat{\rho}_2$ then follows by Lemmas B.7 and B.11.

Lastly, consider $\hat{\alpha}_1$ and $\hat{\alpha}_2$. If Assumptions 1-3 hold with $\Pi^\top\Pi/\sqrt{K} \to \infty$, then by (C.20), (C.26) and Lemma B.7, we have $\widehat{\Phi}_1/\Phi_1 \xrightarrow{p} 1$ and $\widehat{\Phi}_2/\Phi_2 \xrightarrow{p} 1$. Therefore, if the assumptions for d_n in Assumption 4 hold, then by the continuous mapping theorem, we have $\hat{\alpha}_1 \xrightarrow{p} \alpha_1$ and $\hat{\alpha}_2 \xrightarrow{p} \alpha_2$. Alternatively, if Assumption 1-3 hold with $\Pi^\top\Pi/\sqrt{K} = O(1)$, we have $a_2 = 0$, so that $\alpha_2 = 0$ and $\alpha_1 = 1$. By Lemma B.7, we have $\hat{\Sigma}/\Sigma \xrightarrow{p} 1$, and note that $\Sigma/\Gamma_{\tilde{V},\tilde{e}} \to 1$ since $\Pi^\top\Pi/K \to 0$. With $\Gamma_{\tilde{V},\tilde{e}}/K \to \Gamma_{22} > 0$, we can show that $\widehat{\Phi}_2 \rightsquigarrow \bar{\Phi}_2$ for some random variable $\bar{\Phi}_2$ such that $\bar{\Phi}_2 > 0$ with probability one, as in the proof of Step 3 of Lemma B.8. Combining this with the fact that $\widehat{\Phi}_1 \xrightarrow{p} 0$ by (C.20) and Lemma B.7, we have $\hat{\alpha}_2 \xrightarrow{p} 0$ and $\hat{\alpha}_1 \xrightarrow{p} 1$, by the continuous mapping theorem. This concludes the proof.

□

D Proofs of Main Results

D.1 Proof of Proposition 3.1

The proof follows exactly the same lines as in the proof of Lemma 2.2. in Lim et al. (2024) and is thus omitted. □

D.2 Proof of Theorem 4.1

The result follows from Lemma B.9, Lemma B.10 and the Slutsky theorem. □

D.3 Proof of Theorem 4.2

Consider the set $\mathcal{M}'$ of data generating processes m that satisfy the weak convergence result (3.5) pointwise for all $\delta \in \Re$. It is straightforward to see that $\mathcal{M} \subset \mathcal{M}'$. As a result, the test class $\mathfrak{C}$ under consideration is a subset of an augmented class $\mathfrak{C}'$ of ϕ_n satisfying that

$$\lim_{n\to\infty} \mathbb{E}\left[\phi_n\right] \leqslant \alpha \quad \text{for all } m \in \mathcal{M}', \delta = 0, \tag{D.1}$$

$$\liminf_{n\to\infty} \mathbb{E}\left[\phi_n\right] \geqslant \alpha \quad \text{for all } m \in \mathcal{M}', \delta \neq 0. \tag{D.2}$$

We also note that the oracle version of the combination test, ϕ_n^o in (3.6), can be understood as taking the weak convergence result (3.5) as the starting point and is simply the UMPU test in the limiting problem (under known $a_1(\alpha_1), a_2(\alpha_2), \rho_1, \rho_2$), evaluated at the sample analgoues (Wald, LM, and AR statistics). By construction, ϕ_n^o satisfies (D.1) and (D.2), so $\phi_n^o \in \mathfrak{C}'$. Furthermore, by a direct application of Theorem 1 in Müller (2011), it follows that for any $\delta_1 \neq 0$ and any $\phi_n \in \mathfrak{C}'$,

$$\lim_{n\to\infty} \mathbb{E}\left[\phi_n\right] \leqslant \lim_{n\to\infty} \mathbb{E}\left[\phi_n^o\right] \quad \text{for all } m \in \mathcal{M}', \delta = \delta_1,$$

which subsequently implies that for any $\delta_1 \neq 0$ and any $\phi_n \in \mathfrak{C}$,

$$\lim_{n\to\infty} \mathbb{E}\left[\phi_n\right] \leqslant \lim_{n\to\infty} \mathbb{E}\left[\phi_n^o\right] \quad \text{for all } m \in \mathcal{M}, \delta = \delta_1.$$

Lastly, note that $\mathbb{E}\left[\phi_n^*\right]$ is continuous in $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\rho}_1, \hat{\rho}_2)$, which are consistent for $(\alpha_1, \alpha_2, \rho_1, \rho_2)$ under the assumptions of Theorem 4.1 by Lemma B.12. By the continuous mapping theorem, we thus have that, for any δ, $\mathbb{E}\left[\phi_n^*\right] = \mathbb{E}\left[\phi_n^o\right] + o_p(1)$. Therefore, we have ϕ_n^* satisfies (4.1) and (4.2), and thus $\phi_n^* \in \mathfrak{C}$. In addition, for any $\delta_1 \neq 0$, and any $\phi_n \in \mathfrak{C}$,

$$\lim_{n\to\infty} \mathbb{E}\left[\phi_n\right] \leqslant \lim_{n\to\infty} \mathbb{E}\left[\phi_n^o\right] = \lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] \quad \text{for all } m \in \mathcal{M}, \delta = \delta_1.$$

For the second part, it is straightforward to check that $\tilde{\phi}_n \in \mathfrak{C}$. Also, note that the comparison of local asymptotic power of ϕ_n^* and $\tilde{\phi}_n$ can be reduced to the comparison of the non-centrality parameters for ϕ_n^o and $\tilde{\phi}_n$ in their limiting distributions, which are given by $(a^\top V^{-1} a)\delta^2$ and $a_1^2\delta^2$, respectively, where

$$a = \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & \rho_1 & 0 \\ \rho_1 & 1 & \rho_2 \\ 0 & \rho_2 & 1 \end{pmatrix}.$$

Direct calculation yields

$$(a^\top V^{-1} a)\delta^2 - a_1^2\delta^2 = \frac{\delta^2}{1 - \rho_1^2 - \rho_2^2} \left(a_2 - \rho_1 a_1\right)^2 \geqslant 0,$$

and since $\rho_1^2 + \rho_2^2 < 1$ and $\delta \neq 0$, we obtain the desired result. $\square$

D.4 Proof of Theorem 4.3

We shall distinguish between the two cases: (i) $\Pi^\top\Pi/\sqrt{K}$ is diverging and (ii) $\Pi^\top\Pi/\sqrt{K}$ is bounded, and argue along the appropriate subsequence as in Step 2 and 3 in the proof of

Lemma B.8.

For the first case, we have

$$
\begin{aligned}
T(\beta_0) &= \frac{(X^\top z\hat{A}_n z^\top X)^{-1}(X^\top z\hat{A}_n z^\top e)}{\sqrt{\widehat{\Phi}_1}} + \frac{\delta}{\sqrt{\widehat{\Phi}_1}}, \\
LM(\beta_0) &= \frac{X^\top (P-\bar{P})e}{\sqrt{\hat{\Sigma}}} + \frac{\delta(X^\top (P-\bar{P})X)}{\sqrt{\hat{\Sigma}}}, \\
AR &= \frac{\hat{e}^\top (P-\bar{P})\hat{e}}{\sqrt{\hat{\Upsilon}}}
\end{aligned}
$$

under the fixed alternative. On the event $X^\top (P-\bar{P})X > 0$, we can write

$$
\frac{\delta(X^\top (P-\bar{P})X)}{\sqrt{\hat{\Sigma}}} = \frac{\delta}{\sqrt{(X^\top (P-\bar{P})X)^{-1}\hat{\Sigma}(X^\top (P-\bar{P})X)^{-1}}} = \frac{\delta}{\sqrt{\widehat{\Phi}_2}}.
$$

By repeating the proof of Lemma B.9 for the other terms in the above expression beside $\delta/\sqrt{\widehat{\Phi}_1}$ and $\delta/\sqrt{\widehat{\Phi}_2}$ (which do not depend on a_1 and a_2), we can write

$$
\begin{pmatrix} T(\beta_0) \\ LM(\beta_0) \\ AR \end{pmatrix} = \begin{pmatrix} \delta/\sqrt{\widehat{\Phi}_1} \\ \delta/\sqrt{\widehat{\Phi}_2} \\ 0 \end{pmatrix} + R_n,
$$

where $R_n = O_p(1)$. Recall the definition $(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)'$:

$$
\begin{pmatrix} \hat{\omega}_1 \\ \hat{\omega}_2 \\ \hat{\omega}_3 \end{pmatrix} = \frac{1}{\sqrt{\hat{b}_1^2 + \hat{b}_2^2 + \hat{b}_3^2}} \times \begin{pmatrix} 1 & \hat{\rho}_1 & 0 \\ \hat{\rho}_1 & 1 & \hat{\rho}_2 \\ 0 & \hat{\rho}_2 & 1 \end{pmatrix}^{-1/2} \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \end{pmatrix}, \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \end{pmatrix} = \begin{pmatrix} 1 & \hat{\rho}_1 & 0 \\ \hat{\rho}_1 & 1 & \hat{\rho}_2 \\ 0 & \hat{\rho}_2 & 1 \end{pmatrix}^{-1/2} \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ 0 \end{pmatrix},
$$

and note that $(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)' = O_p(1)$. We have

$$
\begin{aligned}
&0 \leqslant 1 - \mathbb{E}\left[\phi_n^*\right] \\
&\quad = \mathbb{P}\left((\hat{\omega}_1 T(\beta_0) + \hat{\omega}_2 LM(\beta_0) + \hat{\omega}_3 AR)^2 < \mathbb{C}_\alpha\right)
\end{aligned}
$$

$$\leqslant \mathbb{P}\left((\hat{\omega}_1 T(\beta_0) + \hat{\omega}_2 LM(\beta_0) + \hat{\omega}_3 AR)^2 < \mathbb{C}_\alpha, X^\top(P-\bar{P})X > 0\right) + \mathbb{P}\left(X^\top(P-\bar{P})X \leqslant 0\right)$$

$$\leqslant \mathbb{P}\left(\left(\delta\sqrt{\frac{1}{\widehat{\Phi}_1}+\frac{1}{\widehat{\Phi}_2}} \times \sqrt{\hat{\alpha}^\top \hat{V}^{-1}\hat{\alpha}} + O_p(1)\right)^2 < \mathbb{C}_\alpha\right) + \mathbb{P}\left(X^\top(P-\bar{P})X \leqslant 0\right),$$

where

$$\hat{\alpha} \equiv \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ 0 \end{pmatrix}, \quad \hat{V} \equiv \begin{pmatrix} 1 & \hat{\rho}_1 & 0 \\ \hat{\rho}_1 & 1 & \hat{\rho}_2 \\ 0 & \hat{\rho}_2 & 1 \end{pmatrix}, \quad \begin{pmatrix} \delta/\sqrt{\widehat{\Phi}_1} \\ \delta/\sqrt{\widehat{\Phi}_2} \\ 0 \end{pmatrix} = \delta\sqrt{\frac{1}{\widehat{\Phi}_1}+\frac{1}{\widehat{\Phi}_2}} \times \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ 0 \end{pmatrix},$$

and we used the fact that $\hat{\omega}_1 \times \hat{\alpha}_1 + \hat{\omega}_2 \times \hat{\alpha}_2 + \hat{\omega}_3 \times 0 = \sqrt{\hat{\alpha}^\top \hat{V}^{-1}\hat{\alpha}}$. In addition, we have $\widehat{\Phi}_1/\Phi_1 \xrightarrow{p} 1$ and $\widehat{\Phi}_2/\Phi_2 \xrightarrow{p} 1$ by (C.20), (C.26) and Lemma B.7, whence $\widehat{\Phi}_1 = o_p(1)$ and $\widehat{\Phi}_2 = o_p(1)$ by (C.21) and (C.25). Therefore, we have

$$\delta\sqrt{\frac{1}{\widehat{\Phi}_1}+\frac{1}{\widehat{\Phi}_2}} \times \sqrt{\hat{\alpha}^\top \hat{V}^{-1}\hat{\alpha}} \xrightarrow{p} \infty(-\infty), \quad \text{for } \delta > 0 \ (\delta < 0),$$

where we use the fact that $\|\hat{\alpha}\|_2 = 1$ by construction, $\hat{V}^{-1} \xrightarrow{p} V^{-1}$ where

$$V \equiv \begin{pmatrix} 1 & \rho_1 & 0 \\ \rho_1 & 1 & \rho_2 \\ 0 & \rho_2 & 1 \end{pmatrix},$$

by Lemma B.12, whence $\hat{\alpha}^\top \hat{V}^{-1}\hat{\alpha}$ can be bounded away from zero with probability approaching one. Together with the fact that $X^\top(P-\bar{P})X > 0$ with probability approaching one by (C.26) and $\Pi^\top\Pi \to \infty$, this implies that $\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] = 1$.

For the second case, by Lemma B.12, we have $\hat{\rho}_1 \xrightarrow{p} \rho_1$ and $\hat{\rho}_2 \xrightarrow{p} \rho_2$, and note that $\rho_1 = 0$ in this case since $\Pi^\top\Pi/K \to 0$. In addition, similar to the proof of Step 3 of Lemma B.8, it can be shown that $\widehat{\Phi}_2 \rightsquigarrow \bar{\Phi}_2$ for some random variable $\bar{\Phi}_2$ such that $\bar{\Phi}_2 > 0$ with probability one, and since $\widehat{\Phi}_1 \xrightarrow{p} 0$ by (C.20), (C.21) and Lemma B.7, it follows that $\hat{\alpha}_1 \xrightarrow{p} 1$ and $\hat{\alpha}_2 \xrightarrow{p} 0$, by the continuous mapping theorem. Therefore, we have $\hat{\omega}_1 \xrightarrow{p} 1$, $\hat{\omega}_2 \xrightarrow{p} 0$ and $\hat{\omega}_3 \xrightarrow{p} 0$. Finally, as in the proof for the first case above, we have $T(\beta_0) = \delta/\sqrt{\widehat{\Phi}_1} + O_p(1)$ where $\widehat{\Phi}_1 = o_p(1)$, $LM(\beta_0) = O_p(1)$ and $AR = O_p(1)$. It follows that $\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] = 1$. This concludes the proof. □

D.5 Proof of Theorem A.1

We shall argue along the appropriate subsequence as in Step 4 of the proof of Lemma B.8. Suppose we are under the local alternative that $\beta - \beta_0 = \delta d_n$. We have

$$
\begin{aligned}
\hat{\rho}_1 &= \frac{1}{\sqrt{\hat{\Psi}\hat{\Sigma}}} \sum_{g\in[G]} \left[\left(\acute{X}_{[g]}^\top \hat{e}_{[g]}\right)\left(\hat{X}_{[g]}^\top \hat{e}_{[g]}\right)\right] \\
&= \sqrt{\frac{\Sigma}{\hat{\Sigma}}} \times \frac{1}{\sqrt{X^\top z \hat{A}_n \hat{\Omega} \hat{A}_n z^\top X}} X^\top z \hat{A}_n \sqrt{\hat{\Omega}} \\
&\quad \times \sqrt{\frac{\Omega}{\hat{\Omega}}} \times \frac{1}{\sqrt{\Omega\Sigma}} \sum_{g\in[G]} \left[\left(z_{[g]}^\top \hat{e}_{[g]}\right)\left(\hat{X}_{[g]}^\top \hat{e}_{[g]}\right)\right],
\end{aligned}
$$

and note the important fact that

$$
\left(\frac{1}{\sqrt{X^\top z \hat{A}_n \hat{\Omega} \hat{A}_n z^\top X}} X^\top z \hat{A}_n \sqrt{\hat{\Omega}}\right)^2 = 1
$$

when $d_z = 1$, and thus

$$
\hat{\rho}_1^2 = \frac{\Sigma}{\hat{\Sigma}} \times \frac{\Omega}{\hat{\Omega}} \times \left(\frac{1}{\sqrt{\Omega\Sigma}} \sum_{g\in[G]} \left[\left(z_{[g]}^\top \hat{e}_{[g]}\right)\left(\hat{X}_{[g]}^\top \hat{e}_{[g]}\right)\right]\right)^2.
$$

By Lemmas B.7 and B.11, we have $\hat{\rho}_1^2 \xrightarrow{p} \bar{\rho}_1^2$. Together with the fact that $\hat{\rho}_2^2 \xrightarrow{p} \rho_2^2$ by Lemmas B.7 and B.11, this implies that $\hat{\rho}_1^2 + \hat{\rho}_2^2 < 1$ with probability approaching one. On that event, direct calculation yields

$$
\hat{\omega}_1 T(\beta_0) + \hat{\omega}_2 LM(\beta_0) + \hat{\omega}_3 AR
$$

$$= \frac{T(\beta_0)\hat{\alpha}_1(1-\hat{\rho}_2^2) - T(\beta_0)\hat{\rho}_1\hat{\alpha}_2 + LM(\beta_0)(\hat{\alpha}_2 - \hat{\rho}_1\hat{\alpha}_1) + AR\hat{\rho}_2(\hat{\rho}_1\hat{\alpha}_1 - \hat{\alpha}_2)}{\sqrt{1-\hat{\rho}_1^2-\hat{\rho}_2^2}\sqrt{\hat{\alpha}_1^2(1-\hat{\rho}_2^2) - 2\hat{\rho}_1\hat{\alpha}_1\hat{\alpha}_2 + \hat{\alpha}_2^2}}.$$

We shall analyze each term in turn. To begin with, we note that, similar to the proof of Step 4 of Lemma B.8, it can be shown that $\widehat{\Phi}_1 \rightsquigarrow \bar{\Phi}_1$ for some random variable $\bar{\Phi}_1$ such that $\bar{\Phi}_1 > 0$ with probability one, and since $\widehat{\Phi}_2 \xrightarrow{p} 0$ by (C.25), (C.26) and Lemma B.7, it follows that $\hat{\alpha}_1 \xrightarrow{p} 0$ and $\hat{\alpha}_2 \xrightarrow{p} 1$, by the continuous mapping theorem. Note also that

$$\begin{aligned} T(\beta_0) &= \frac{(X^\top z\hat{A}_n z^\top X)^{-1}}{\sqrt{(X^\top z\hat{A}_n z^\top X)^{-2}}} \times \frac{1}{\sqrt{X^\top z\hat{A}_n\hat{\Omega}\hat{A}_n z^\top X}} X^\top z\hat{A}_n\sqrt{\hat{\Omega}} \\ &\quad \times \sqrt{\frac{\Omega}{\hat{\Omega}}} \times \frac{1}{\sqrt{\Omega}} \sum_{g\in[G]} z_{[g]}^\top \tilde{e}_{[g]} + \frac{d_n}{\sqrt{\widehat{\Phi}_1}}\delta \\ &= O_p(1), \end{aligned}$$

and thus

$$T(\beta_0)\hat{\alpha}_1(1-\hat{\rho}_2^2) = o_P(1).$$

Next, we note that

$$\begin{aligned} T(\beta_0)\hat{\rho}_1 = & \frac{(X^\top z\hat{A}_n z^\top X)^{-1}}{\sqrt{(X^\top z\hat{A}_n z^\top X)^{-2}}} \times \sqrt{\frac{\Sigma}{\hat{\Sigma}}} \times \frac{\Omega}{\hat{\Omega}} \\ & \times \frac{1}{\sqrt{\Omega}} \sum_{g\in[G]} z_{[g]}^\top \tilde{e}_{[g]} \times \frac{1}{\sqrt{\Omega\Sigma}} \sum_{g\in[G]} \left[\left(z_{[g]}^\top \hat{e}_{[g]}\right) \left(\hat{X}_{[g]}^\top \hat{e}_{[g]}\right) \right] + o_P(1). \end{aligned}$$

In addition, we can show that

$$\frac{1}{\sqrt{n}} z^\top X \rightsquigarrow \mathcal{N}\left(\pi, \Omega_z^{\tilde{V},\tilde{V}}\right),$$

which, combining with the fact that $\hat{A}_n/\lambda_n \xrightarrow{p} A$, implies that

$$\frac{(X^\top z\hat{A}_n z^\top X)^{-1}}{\sqrt{(X^\top z\hat{A}_n z^\top X)^{-2}}} = 1 + o_P(1).$$

If follows that, by Lemmas B.7 and B.11,

$$T(\beta_0)\hat{\rho}_1\hat{\alpha}_2 = \bar{\rho}_1 \times \frac{1}{\sqrt{\Omega}} \sum_{g\in[G]} z_{[g]}^\top \tilde{e}_{[g]} + o_P(1).$$

Next, we note that

$$\begin{aligned}
LM(\beta_0) &= \frac{X^\top(P-\bar{P})e}{\sqrt{\hat{\Sigma}}} + a\delta + o_P(1) \\
&= \frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]} \hat{\Pi}_{[g]}^\top \tilde{e}_{[g]} + \sum_{g,h\in[G]^2, g\neq h} \tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right) + a\delta + o_P(1),
\end{aligned}$$

where the first equality is by (C.26) and Lemma B.7, and the second equality is by Lemmas B.3 and B.7. It follows that

$$LM(\beta_0)(\hat{\alpha}_2 - \hat{\rho}_1\hat{\alpha}_1) = \frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]} \hat{\Pi}_{[g]}^\top \tilde{e}_{[g]} + \sum_{g,h\in[G]^2, g\neq h} \tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right) + a\delta + o_P(1).$$

Next, we note that

$$\begin{aligned}
AR &= \frac{1}{\sqrt{\widehat{\Upsilon}}}\left(\hat{e}^\top(P-\bar{P})\hat{e} - e^\top(P-\bar{P})e\right) + \frac{1}{\sqrt{\widehat{\Upsilon}}}\left(e^\top(P-\bar{P})e - \tilde{e}^\top(P-\bar{P})\tilde{e}\right) + \frac{1}{\sqrt{\widehat{\Upsilon}}}\tilde{e}^\top(P-\bar{P})\tilde{e} \\
&= \frac{1}{\sqrt{\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]} + o_P(1).
\end{aligned}$$

by (C.22), (C.24), and Lemmas B.3 and B.7. It follows that

$$AR\hat{\rho}_2(\hat{\rho}_1\hat{\alpha}_1 - \hat{\alpha}_2) = -\rho_2 \times \frac{1}{\sqrt{\Upsilon}} \sum_{g,h\in[G]^2, g\neq h} \tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]} + o_P(1)$$

by Lemma B.11. Finally, we have

$$1-\hat{\rho}_1^2-\hat{\rho}_2^2 \xrightarrow{p} 1-\bar{\rho}_1^2-\rho_2^2,$$
$$\hat{\alpha}_1^2(1-\hat{\rho}_2^2)-2\hat{\rho}_1\hat{\alpha}_1\hat{\alpha}_2+\hat{\alpha}_2^2 \xrightarrow{p} 1.$$

Combining all the results, we have

$$\begin{aligned}
&\hat{\omega}_1 T(\beta_0)+\hat{\omega}_2 LM(\beta_0)+\hat{\omega}_3 AR \\
=&\frac{-\bar{\rho}_1}{\sqrt{1-\bar{\rho}_1^2-\rho_2^2}}\times\frac{1}{\sqrt{\Omega}}\sum_{g\in[G]} z_{[g]}^\top \tilde{e}_{[g]} \\
&+\frac{1}{\sqrt{1-\bar{\rho}_1^2-\rho_2^2}}\times\left(\frac{1}{\sqrt{\Sigma}}\left(\sum_{g\in[G]}\hat{\Pi}_{[g]}^\top\tilde{e}_{[g]}+\sum_{g,h\in[G]^2,g\neq h}\tilde{V}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}\right)+a\delta\right) \\
&+\frac{-\rho_2}{\sqrt{1-\bar{\rho}_1^2-\rho_2^2}}\times\frac{1}{\sqrt{\Upsilon}}\sum_{g,h\in[G]^2,g\neq h}\tilde{e}_{[g]}^\top P_{[g,h]}\tilde{e}_{[h]}+o_P(1).
\end{aligned}$$

Furthermore, we note that in the proof of Lemma B.10, we only require $\acute{\Pi}$ to satisfy

$$\frac{1}{\Psi}\sum_{g\in[G]}\acute{\Pi}_{[g]}^\top\acute{\Pi}_{[g]}=O(1),\quad \frac{1}{\Psi}\max_{g\in[G]}\acute{\Pi}_{[g]}^\top\acute{\Pi}_{[g]}=o(1),$$

which is guaranteed by Assumptions 1 and 2. We also have

$$\frac{1}{\Omega}\sum_{g\in[G]} z_{[g]}^\top z_{[g]}=O(1),\quad \frac{1}{\Omega}\max_{g\in[G]} z_{[g]}^\top z_{[g]}=o(1),$$

by Assumption 1 alone. Therefore, we can replace $\frac{1}{\sqrt{\Psi}}\sum_{g=1}^G\acute{\Pi}_{[g]}^\top\tilde{e}_{[g]}$ with $\frac{1}{\sqrt{\Omega}}\sum_{g\in[G]} z_{[g]}^\top\tilde{e}_{[g]}$ in the proof of Lemma B.10. Note that in this way, Assumption 2 is no longer needed. We thus can follow the same argument in the proof of Lemma B.10 and obtain that

$$\begin{pmatrix} \frac{1}{\sqrt{\Omega}} \sum_{g \in [G]} z_{[g]}^\top \tilde{e}_{[g]} \\ \frac{1}{\sqrt{\Sigma}} \left(\sum_{g \in [G]} \hat{\Pi}_{[g]}^\top \tilde{e}_{[g]} + \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]} \right) + a\delta \\ \frac{1}{\sqrt{\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]} \end{pmatrix} \rightsquigarrow \mathcal{N} \left(\begin{pmatrix} 0 \\ a\delta \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \bar{\rho}_1 & 0 \\ \bar{\rho}_1 & 1 & \rho_2 \\ 0 & \rho_2 & 1 \end{pmatrix} \right).$$

The desired result then follows.

Next, suppose we are under the fixed alternative. From the proof above, we have

$$\begin{aligned} & \hat{\omega}_1 T(\beta_0) + \hat{\omega}_2 LM(\beta_0) + \hat{\omega}_3 AR \\ = & \frac{(1 + o_P(1))}{\sqrt{1 - \rho_1^2 - \rho_2^2}} LM(\beta_0) + O_P(1) \\ = & \frac{(1 + o_P(1))}{\sqrt{1 - \rho_1^2 - \rho_2^2}} \times \left(\frac{\delta}{\sqrt{\widehat{\Phi}_2}} \times (1 + o_P(1)) + O_P(1) \right) + O_P(1) \end{aligned}$$

by (C.26) and Lemma B.7, and note that $\widehat{\Phi}_2 = o_P(1)$ by (C.25) and Lemma B.7. If follows that

$$(\hat{\omega}_1 T(\beta_0) + \hat{\omega}_2 LM(\beta_0) + \hat{\omega}_3 AR)^2 \xrightarrow{p} \infty,$$

and the desired result follows. This concludes the proof. □

D.6 Proof of Theorem A.2

Under the local alternative, similar to the proof of Theorem A.1, we have

$$\begin{aligned} T(\beta_0) = & \frac{(X^\top z \hat{A}_n z^\top X)^{-1}}{\sqrt{(X^\top z \hat{A}_n z^\top X)^{-2}}} \times \frac{1}{\sqrt{X^\top z \hat{A}_n \hat{\Omega} \hat{A}_n z^\top X}} X^\top z \hat{A}_n \hat{\Omega}^{1/2} \\ & \times \hat{\Omega}^{-1/2} \Omega^{1/2} \times \Omega^{-1/2} \sum_{g \in [G]} z_{[g]}^\top \tilde{e}_{[g]} + \frac{d_n}{\sqrt{\widehat{\Phi}_1}} \delta = O_p(1), \end{aligned}$$

$$LM(\beta_0) = \frac{1}{\sqrt{\Sigma}} \left(\sum_{g \in [G]} \hat{\Pi}_{[g]}^\top \tilde{e}_{[g]} + \sum_{g,h \in [G]^2, g \neq h} \tilde{V}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]} \right) + a\delta + o_p(1),$$
$$AR = \frac{1}{\sqrt{\Upsilon}} \sum_{g,h \in [G]^2, g \neq h} \tilde{e}_{[g]}^\top P_{[g,h]} \tilde{e}_{[h]} + o_p(1).$$

By repeating the proof of Lemma B.10 and ignoring $\frac{1}{\sqrt{\Psi}} \sum_{g=1}^{G} \acute{\Pi}_{[g]}^\top \tilde{e}_{[g]}$ (and thus Assumption 2 is not needed), we have

$$\begin{pmatrix} LM(\beta_0) \\ AR \end{pmatrix} \rightsquigarrow \mathcal{N}\left(\begin{pmatrix} a\delta \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right),$$

by the Slutsky theorem. For $\hat{\rho}_1$, we have

$$\begin{aligned} \hat{\rho}_1 &= \frac{1}{\sqrt{\hat{\Psi}\hat{\Sigma}}} \sum_{g \in [G]} \left[\left(\acute{X}_{[g]}^\top \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^\top \hat{e}_{[g]} \right) \right] \\ &= \sqrt{\frac{\Sigma}{\hat{\Sigma}}} \times \frac{1}{\sqrt{X^\top z \hat{A}_n \hat{\Omega} \hat{A}_n z^\top X}} X^\top z \hat{A}_n \hat{\Omega}^{1/2} \\ &\times \hat{\Omega}^{-1/2} \Omega^{1/2} \times \frac{1}{\sqrt{\Sigma}} \Omega^{-1/2} \sum_{g \in [G]} \left[\left(z_{[g]}^\top \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^\top \hat{e}_{[g]} \right) \right], \end{aligned}$$

where

$$\frac{1}{\sqrt{X^\top z \hat{A}_n \hat{\Omega} \hat{A}_n z^\top X}} X^\top z \hat{A}_n \hat{\Omega}^{1/2} = O_P(1),$$

and by Lemma B.11,

$$\frac{1}{\sqrt{\Sigma}} \Omega^{-1/2} \sum_{g \in [G]} \left[\left(z_{[g]}^\top \hat{e}_{[g]} \right) \left(\hat{X}_{[g]}^\top \hat{e}_{[g]} \right) \right] = \frac{1}{\sqrt{\Sigma}} \Omega^{-1/2} \sum_{g \in [G]} \mathbb{E} \left[\left(z_{[g]}^\top \tilde{e}_{[g]} \right) \left(\hat{\Pi}_{[g]}^\top \tilde{e}_{[g]} \right) \right] + o_P(1).$$

Further note that the first term on the right-hand side of the above display is $o(1)$ since $\Pi^\top \Pi / K \to 0$, so that $\hat{\rho}_1 \xrightarrow{p} 0$ by Lemma B.7. In addition, we have $\hat{\rho}_2 \xrightarrow{p} \rho$ by Lemmas B.7 and B.11. Finally, similar to the proof of Step 4 of Lemma B.8, it can be shown that $\widehat{\Phi}_1 \rightsquigarrow \bar{\Phi}_1$ for some random variable $\bar{\Phi}_1$ such that $\bar{\Phi}_1 > 0$ with probability one, and since $\widehat{\Phi}_2 \xrightarrow{p} 0$ by (C.25), (C.26) and Lemma B.7, it follows that $\hat{\alpha}_1 \xrightarrow{p} 0$ and $\hat{\alpha}_2 \xrightarrow{p} 1$, by the continuous mapping theorem. This implies that $\hat{\omega}_1 \xrightarrow{p} 0$, $\omega_2 \xrightarrow{p} 1/\sqrt{1-\rho^2}$ and $\omega_3 \xrightarrow{p} -\rho/\sqrt{1-\rho^2}$, whence

$$\begin{aligned}
&\hat{\omega}_1 T(\beta_0) + \hat{\omega}_2 LM(\beta_0) + \hat{\omega}_3 AR \\
&= \frac{1}{\sqrt{1-\rho^2}} LM(\beta_0) - \frac{\rho}{\sqrt{1-\rho^2}} AR + o_p(1) \\
&\rightsquigarrow \frac{1}{\sqrt{1-\rho^2}} \mathcal{N}_1 - \frac{\rho}{\sqrt{1-\rho^2}} \mathcal{N}_2,
\end{aligned}$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ are defined in Theorem A.2.

Under the fixed alternative, similar to the proof of Theorem A.1, we have $\lim_{n\to\infty} \mathbb{E}\left[\phi_n^*\right] = 1$. This concludes the proof. □

E Additional Simulations

In this section, we present some additional simulation results to illustrate the effect of weak low-dimensional IVs. We set $\psi = 30$ so that the identification strength of many IVs remains relatively strong, and $\phi = 0$ so that the identification strength of the one-dimensional IV is rather weak. Figure 1 displays the power curves for $K = 100$ and $K = 500$, respectively, which can be regarded as extensions of Panels A and B of Figure 3 in the main text. Overall, the dominant performance of our combination test remains robust to different strengths of one-dimensional IV, and the Wald test based on one-dimensional IV provides a nontrivial gain in power (as seen from the noticeable gaps between the power curves of ϕ_n^* and the LM test). This gain arises from its correlation with the LM statistic, in line with the theoretical result stated in Theorem A.1.

F Additional Empirical Applications

In this section, we consider the return to education application, using the dataset of Angrist and Krueger (1991). In this application, the outcome variable is the log weekly wages and the endogenous variable is the years of schooling. We follow Mikusheva and Sun (2022) and Lim et al. (2024) to consider two specifications with 180 and 1530 instruments. The set of 180 instruments consists of 30 quarter and year of birth interactions (QoB–YoB) and 150 quarter and place of birth interactions (QoB-PoB). The set of 1530 instruments includes all interactions among QoB-YoB-PoB. The quantitative implications obtained from Table 1 and Figure 2 are in line with the discussion in Section 2 of the main text. However, we reiterate that, although the low-dimensional IVs (and therefore $\hat{\beta}_1$ and the Wald confidence interval) are identical across the two specifications by construction, we are not attempting to obtain improved inference based on pooling statistics across specifications, and our theoretical results do not justify such an approach.

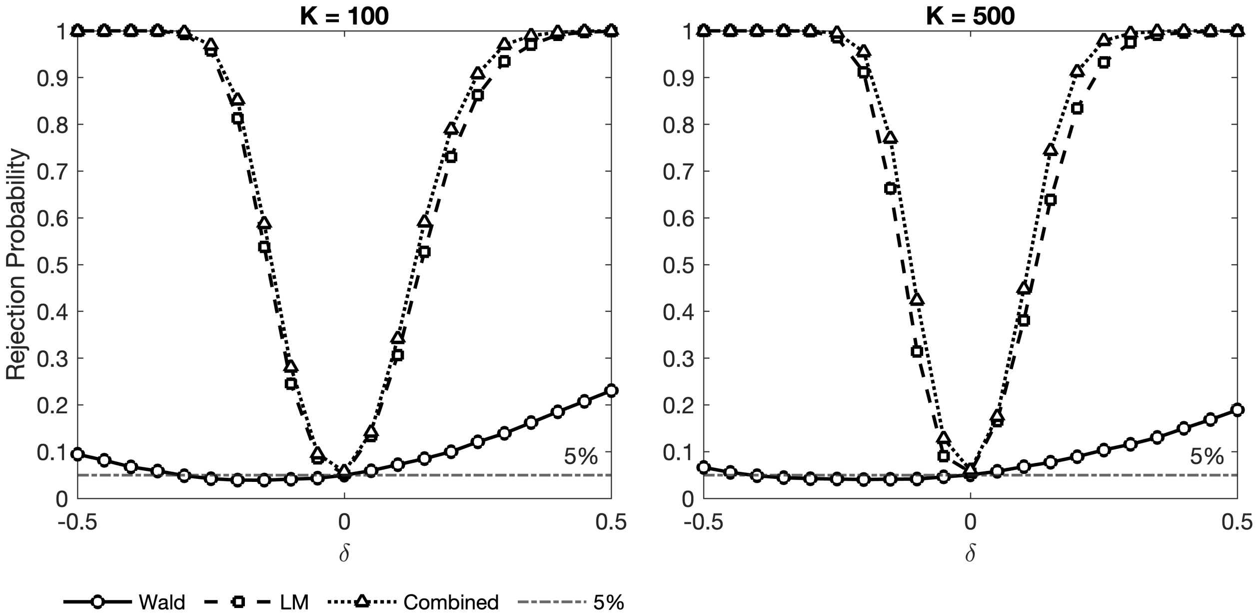

Figure 1: Power Curve of the combination, Wald, and jacknife LM tests.

Notes: This figure displays the power curves for our combination test ϕ_n^* along with those for the component Wald and jacknife LM tests, at different values of K (the dimension of the many IVs), $\psi = 30$ (the identification strength of the many IVs is relatively strong), and $\phi = 0$ (the identification strength of the one-dimensional IV is weak). The horizontal axis represents the deviations in the parameter of interest from the maintained hypothesis, that is, we are interested in testing $\mathcal{H}_0 : \beta = \beta_0$ against $\mathcal{H}_1 : \beta \neq \beta_0$, and $\delta = \beta - \beta_0$. See Section 5 in the main text for a detailed description of the simulation setup. All results are based on $5,000$ simulations.

	$K = 180$	$K = 1530$
$\hat{\rho}_1$	0.489	0.236
$\hat{\rho}_2$	-0.170	-0.200
$\hat{\sigma}(\hat{\beta}_1)/\hat{\sigma}(\hat{\beta}_2)$	1.191	0.807
$\hat{\beta}_1$	0.098	0.098
Wald CI	(0.059, 0.138)	(0.059, 0.138)
$\hat{\beta}_2$	0.099	0.084
LM CI	(0.066, 0.132)	(0.035, 0.133)
$\hat{\beta}^*$	0.097	0.093
Comb. CI	(0.066, 0.127)	(0.059, 0.126)

Table 1: Point estimates and confidence intervals: returns to education.

Notes: This table reports the estimation and inference results for the return to education example using the Angrist and Krueger (1991) dataset, shown separately for specifications with $K = 180$ instruments and $K = 1530$ instruments. The IV set in the column labeled "$K = 180$" consists of 30 quarter and year of birth interactions (QoB–YoB) and 150 quarter and place of birth interactions (QoB–PoB), while the IV set in the column with "$K = 1530$" includes full set of interactions among QoB–YoB–PoB. See Appendix D in Lim et al. (2024) for a more detailed description of data and this empirical application. The point estimates are obtained from the standard two-stage least squares (TSLS) estimator with the three-dimensional QoB instruments, $\hat{\beta}_1$, and, in addition, from the leave-one-out estimator, $\hat{\beta}_2$, which makes use of all base IVs. Wald CI and LM CI denote the confidence intervals based on $\hat{\beta}_1$ and $\hat{\beta}_2$, respectively. The estimator $\hat{\beta}^*$ is the combined estimator for β, defined in Section 4.3 of the main text. It is essentially the midpoint of the confidence interval in (4.4), which is obtained from our combination test and labeled as "Comb. CI" in the table. In addition, $\hat{\rho}_1$ and $\hat{\rho}_2$ denote estimates of the asymptotic correlation between the Wald and LM statistics, and between the LM and AR statistics, respectively. Finally, $\hat{\sigma}(\hat{\beta}_1)/\hat{\sigma}(\hat{\beta}_2)$ denotes the ratio of standard errors of $\hat{\beta}_1$ and $\hat{\beta}_2$. All displayed numbers are rounded to three decimal places.

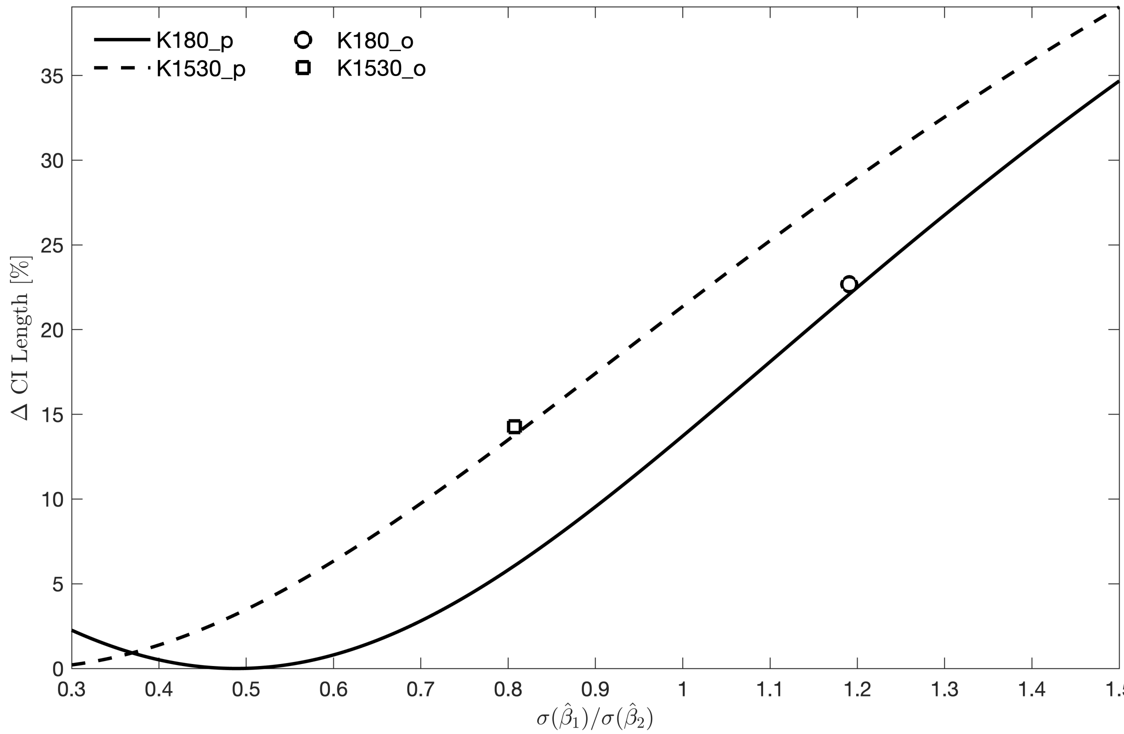

Figure 2: Realized percentage reduction in confidence interval length: returns to education

Notes: This figure shows, for each specification in the returns to education example, the observed percentage decrease in confidence interval length (Combined CI versus Wald CI, as in Table 1, and indicated by "o" in figure legends) plotted as a point against the standard error ratio ($\hat{\sigma}(\hat{\beta}_1)/\hat{\sigma}(\hat{\beta}_2)$ in Table 1). Also shown is the theoretical lower bound for the reduction (indicated by "p" in figure legends), analogous to Figure 1 in the main text, but now computed using the specification-specific estimate $\hat{\rho}_1$, as reported in Table 1. Here, "K180" refers to the specifications with $K = 180$ instruments, and "K1530" refers to the specifications with $K = 1530$ instruments. The horizontal axis is the ratio of standard deviations (errors) of $\hat{\beta}_1$ and $\hat{\beta}_2$. The vertical axis is the reduction in the length of confidence interval in percentage points. As a final remark, note that the actual numerical values of the relevant quantities in Table 1, rather than the rounded values shown there, are used to produce Figure 2.

References

ANDREWS, I., J. H. STOCK, AND L. SUN (2019): "Weak instruments in instrumental variables regression: Theory and practice," *Annual Review of Economics*, 11, 727–753.

ANGRIST, J. D. AND A. B. KRUEGER (1991): "Does Compulsory School Attendance Affect Schooling and Earning?" *Quarterly Journal of Economics*, 106, 979–1014.

CHAO, J. C., N. R. SWANSON, J. A. HAUSMAN, W. K. NEWEY, AND T. WOUTERSEN (2012): "Asymptotic Distribution Of JIVE In A Heteroskedastic IV Regression With Many Instruments," *Econometric Theory*, 28, 42–86.

DJOGBENOU, A. A., J. G. MACKINNON, AND M. O. NIELSEN (2019): "Asymptotic theory and wild bootstrap inference with clustered errors," *Journal of Econometrics*, 212, 393–412.

HALL, P. AND C. C. HEYDE (1980): *Martingale limit theory and its application*, Academic Press.

HANSEN, B. E. AND S. LEE (2019): "Asymptotic theory for clustered samples," *Journal of Econometrics*, 210, 268–290.

LEE, D. S., J. MCCRARY, M. J. MOREIRA, AND J. PORTER (2022): "Valid t-ratio Inference for IV," *American Economic Review*, 112, 3260–3290.

LIM, D., W. WANG, AND Y. ZHANG (2024): "A conditional linear combination test with many weak instruments," *Journal of Econometrics*, 238, 105602.

MAGNUS, J. R. AND H. NEUDECKER (2019): *Matrix differential calculus with applications in statistics and econometrics*, John Wiley & Sons.

MIKUSHEVA, A. AND L. SUN (2022): "Inference with many weak instruments," *Review of Economic Studies*, 89, 2663–2686.

Müller, U. K. (2011): "Efficient tests under a weak convergence assumption," *Econometrica*, 79, 395–435.

Staiger, D. and J. Stock (1997): "Instrumental variables regression with weak instruments," *Econometrica*, 65, 557–586.

Young, A. (2022): "Consistency without inference: Instrumental variables in practical application," *European Economic Review*, 147, 104112.