A Conditionally Studentized Test for High-Dimensional Parametric Regression via Sample Splitting

Feng Liang[1], Chuhan Wang[1], Jiaqi Huang[2], Lixing Zhu[1]

1 Department of Statistics, Beijing Normal University at Zhuhai, China

2 School of Statistics, University of Michigan, Ann Arbor, Michigan

Supplementary Material

S1 Regularity Conditions

For the convenience of theoretical proofs, we repeat the regularity conditions here. We add some explanations for our conditions from Remarks 2 to 3. In these conditions, $\|\cdot\|$ means the L_2 norm of any vector or matrix.

Condition 1. *There exists a unique minimizer $\boldsymbol{\theta}^* \in \mathbb{R}^p$ of the squared loss function, lying in the interior of the compact parameter set $\boldsymbol{\Theta}$.*

Condition 2. *Let $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_p)$. The regression function admits third*

derivatives concerning $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. *For* $\forall j, k = 1, 2, \cdots, p$, *define*

$$\dot{g}(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{\partial g(\boldsymbol{\theta}, \boldsymbol{x})}{\partial \boldsymbol{\theta}}, \quad \ddot{g}(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{\partial^2 g(\boldsymbol{\theta}, \boldsymbol{x})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top},$$

$$\dot{g}_j(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{\partial g(\boldsymbol{\theta}, \boldsymbol{x})}{\partial \theta_j} \quad \text{and} \quad \ddot{g}_{jk}(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{\partial g(\boldsymbol{\theta}, \boldsymbol{x})}{\partial \theta_j \partial \theta_k}.$$

Let $U(\boldsymbol{\theta}^*)$ *be a subset consists of all* $\boldsymbol{\theta}$ *satisfying* $\|\boldsymbol{\theta} - \boldsymbol{\theta}^*\| \leq r_0$ *in the interior of* $\boldsymbol{\Theta}$, *where* r_0 *is a positive constant. Further, for* $\forall j, k = 1, 2, \cdots, p$ *and* $\forall \boldsymbol{\theta} \in U(\boldsymbol{\theta}^*)$, *there exists a function* $F(\cdot)$ *such that for any* $\boldsymbol{x}$,

$$|g(\boldsymbol{\theta}, \boldsymbol{x})| \leq F(\boldsymbol{x}), |\dot{g}_j(\boldsymbol{\theta}, \boldsymbol{x})| \leq F(\boldsymbol{x}), |\ddot{g}_{jk}(\boldsymbol{\theta}, \boldsymbol{x})| \leq F(\boldsymbol{x})$$

with $E\left\{F^4(\boldsymbol{X})\right\} = O(1)$. *The fourth moment* $E(\varepsilon^4)$ *of* ε *is bounded for the regression model.*

For $j = 1, 2, \cdots, p$, define $\psi_j(\boldsymbol{\theta}, \boldsymbol{x}) = \{m(\boldsymbol{x}) - g(\boldsymbol{\theta}, \boldsymbol{x})\} \dot{g}_j(\boldsymbol{\theta}, \boldsymbol{x})$, $P\ddot{\psi}_{j\boldsymbol{\theta}} = E\left\{\frac{\partial \psi_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}(\boldsymbol{\theta}, \boldsymbol{X})\right\}$, and $\lambda_i\left(P\ddot{\psi}_{j\boldsymbol{\theta}}\right)$ be the i-th eigenvalue of $P\ddot{\psi}_{j\boldsymbol{\theta}}$.

Condition 3. *For any* $\boldsymbol{\theta} \in U(\boldsymbol{\theta}^*)$, $\max_{1 \leq i,j \leq p} \lambda_i\left(P\ddot{\psi}_{j\boldsymbol{\theta}}\right) \leq c$, *where* c *is a positive constant free from* n *and* p.

For any matrix $\boldsymbol{A}$, let $\lambda_{\min}(\boldsymbol{A})$ and $\lambda_{\max}(\boldsymbol{A})$ be the smallest and the largest eigenvalue of $\boldsymbol{A}$, respectively.

Condition 4. *Define*

$$\boldsymbol{\Sigma} = E\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X})^\top\right\} - E\left[\{m(\boldsymbol{X}) - g(\boldsymbol{\theta}^*, \boldsymbol{X})\} \ddot{g}(\boldsymbol{\theta}^*, \boldsymbol{X})\right].$$

There exist two constants a and b unrelated to n and p, such that

$$0 < a \leq \lambda_{\min}(\boldsymbol{\Sigma}) \leq \lambda_{\max}(\boldsymbol{\Sigma}) \leq b < \infty.$$

Besides, define

$$\boldsymbol{\Sigma}_* = E\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X})\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X})^\top\right\} \quad \textit{and} \quad \boldsymbol{\Sigma}_\varepsilon = E\left\{\varepsilon^2\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X})\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X})^\top\right\},$$

where ε is the error term corresponding to $\boldsymbol{X}$ in our regression model. There exist four constants a^*, b^*, a_ε and b_ε unrelated to n and p such that

$$0 < a^* \leq \lambda_{\min}(\boldsymbol{\Sigma}_*) \leq \lambda_{\max}(\boldsymbol{\Sigma}_*) \leq b^* < \infty,$$

and

$$0 < a_\varepsilon \leq \lambda_{\min}(\boldsymbol{\Sigma}_\varepsilon) \leq \lambda_{\max}(\boldsymbol{\Sigma}_\varepsilon) \leq b_\varepsilon < \infty.$$

Finally, define $\boldsymbol{\Sigma}^*_{cov} = E\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top w_{12}\right\}$. Assume $\|\boldsymbol{\Sigma}^*_{cov}\| \leq b_{cov}$, where b_{cov} is a constant free from n and p.

Condition 5. *Define*

$$\ddot{\psi}_{jkl}(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{\partial^2\psi_j}{\partial\theta_k\partial\theta_l}(\boldsymbol{\theta}, \boldsymbol{x}) \quad \textit{and} \quad g^{(3)}_{jkl}(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{\partial^3 g}{\partial\theta_j\theta_k\partial\theta_l}(\boldsymbol{\theta}, \boldsymbol{x}).$$

There exist two measurable functions $F_1(\cdot)$ and $F_2(\cdot)$ with $E\{F_1^4(\boldsymbol{X})\} = O(1)$ and $E\{F_2^8(\boldsymbol{X})\} = O(1)$, satisfying

$$\left|\ddot{\psi}_{jkl}(\boldsymbol{\theta}_1, \boldsymbol{x}) - \ddot{\psi}_{jkl}(\boldsymbol{\theta}_2, \boldsymbol{x})\right| \leq \sqrt{p}\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|F_1(\boldsymbol{x})$$

and

$$\left|g^{(3)}_{jkl}(\boldsymbol{\theta}_1, \boldsymbol{x}) - g^{(3)}_{jkl}(\boldsymbol{\theta}_2, \boldsymbol{x})\right| \leq \sqrt{p}\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|F_2(\boldsymbol{x}),$$

where $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in U(\boldsymbol{\theta}^*)$.

Condition 6. *There exists a measurable function* $L(\cdot)$ *with* $E\{L^4(\boldsymbol{X})\} = O(1)$*, satisfying* $\|\dot{g}(\boldsymbol{\theta}_1, \boldsymbol{x}) - \dot{g}(\boldsymbol{\theta}_2, \boldsymbol{x})\| \leq \sqrt{p}\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|L(\boldsymbol{x})$*, where* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in U(\boldsymbol{\theta}^*)$.

Condition 7. $\inf_{\boldsymbol{x}_i, \boldsymbol{x}_j} W_n(\boldsymbol{x}_i, \boldsymbol{x}_j) \geq 0$ *and* $E\{W_n^4(\boldsymbol{X}_1, \boldsymbol{X}_2)\} = E(w_{12}^4) \leq C_1$ *for a positive constant* C_1.

Remark 1. *In fact, under Conditions 4 and 7, we have* $\|E\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\}\| \leq \sqrt{b^*}$ *and* $\|E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)w_{12}\}\| \leq \sqrt{b^* E_{\boldsymbol{X}_1}(w_{12}^2)}$*, where* $E_{\boldsymbol{X}_1} = E(\cdot|\boldsymbol{X}_1)$ *means given* $\boldsymbol{X}_1$*, the conditional expectation over other random variable(s). This is because*

$$
\begin{aligned}
&\|E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12}\}\|^2 \\
=& \sup_{\|\boldsymbol{a}\|=1} \left[\boldsymbol{a}^\top E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)w_{12}\}\right]^2 \\
=& \sup_{\|\boldsymbol{a}\|=1} E_{\boldsymbol{X}_1}^2\left\{\boldsymbol{a}^\top \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)w_{12}\right\} \\
\leq& \sup_{\|\boldsymbol{a}\|=1} E_{\boldsymbol{X}_1}\left\{|\boldsymbol{a}^\top \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)|^2\right\} E_{\boldsymbol{X}_1}\left(w_{12}^2\right) \\
=& \sup_{\|\boldsymbol{a}\|=1} E_{\boldsymbol{X}_1}\left\{\boldsymbol{a}^\top \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top \boldsymbol{a}\right\} E_{\boldsymbol{X}_1}\left(w_{12}^2\right) \\
=& \|E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top\right\}\| E_{\boldsymbol{X}_1}\left(w_{12}^2\right) \\
=& \|E\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top\right\}\| E_{\boldsymbol{X}_1}\left(w_{12}^2\right) = \|\boldsymbol{\Sigma}_*\| E_{\boldsymbol{X}_1}\left(w_{12}^2\right) \leq b^* E_{\boldsymbol{X}_1}\left(w_{12}^2\right),
\end{aligned}
$$

and we can prove $\|E\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\}\| \leq \sqrt{b^*}$ *if we replace* w_{12} *by* 1. *Since*

$E\left(w_{12}^4\right) \leq C_1$, *we also have*

$$\begin{aligned} & E\left[\left\|E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12}\right\}\right\|^4\right] \\ \leq & E\left\{b^{*2} E_{\boldsymbol{X}_1}^2\left(w_{12}^2\right)\right\} \leq b^{*2} E\left(w_{12}^4\right) \leq b^{*2} C_1=: C_2 . \end{aligned}$$

Condition 8. $\max_{1 \leq k, l \leq p} E\left\{\varepsilon_1^4 \dot{g}_k^2(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \dot{g}_l^2(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\right\}<\infty$.

Remark 2. *If* ε *and* $\boldsymbol{X}$ *are independent, Condition 8 can be yield from* $\left|\dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\right| \leq F(\boldsymbol{X}_1)$, $E\left\{F^4(\boldsymbol{X}_1)\right\}=O(1)$ *in Condition 2 and* $E(\varepsilon^4)<\infty$. *If* ε *and* $\boldsymbol{X}$ *are dependent, a sufficient condition for Condition 8 is* $E(\varepsilon^8)<\infty$ *and* $\max_{1 \leq k \leq p} E\left\{\dot{g}_k^4(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\right\}<\infty$, *based on Cauchy-Schwarz inequality. Note that* $\max_{1 \leq k \leq p} E\left\{\dot{g}_k^4(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\right\}<\infty$ *is a direct result of* $E\left\{F^4(\boldsymbol{X}_1)\right\}=O(1)$ *under Condition 2.*

Recall

$$\begin{aligned} {w'}_{ij} = & w_{ij}-\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^{\top} \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{i2}\right\} \\ & -\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^{\top} \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{1 j}\right\} \\ & +\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j), \end{aligned} \tag{S1.1}$$

which is defined before Condition 1.9 in the main context; we repeat it here for convenience.

Condition 9. $E\left(\varepsilon_1^2 \varepsilon_2^2 {w'}_{12}^2\right) \geq C_3$, *and* $E\left(\varepsilon_1^4 \varepsilon_2^4 {w'}_{12}^4\right) \leq C_4$, *where* C_3 *and*

C_4 are two positive constants. Besides,

$$E\left[\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)\right\}^4\right] = O(1).$$

Remark 3. *For Condition 9, we give an example. Assume $\boldsymbol{X}_1 \sim \mathbf{N}(\mathbf{0}, \mathbf{I}_q)$, $\varepsilon_1 \sim \mathbf{N}(0,1)$, $\boldsymbol{X}_1 = (X_{11}, \cdots, X_{1q})^\top$, $Y_1 = \boldsymbol{X}_1^\top \boldsymbol{\theta}_0 + \varepsilon_1$, $p = q$, $w_{12} = \exp\left\{-\frac{1}{p}\sum_{k=1}^p (X_{1k} - X_{2k})^2)\right\}$. Then $\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) = \boldsymbol{X}_1^\top$, $\boldsymbol{\Sigma} = \mathbf{I}_p$, and*

$$E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12}\right\} = E_{\boldsymbol{X}_1}\left[\boldsymbol{X}_2^\top \exp\left\{-\frac{1}{p}\sum_{k=1}^p (X_{1k} - X_{2k})^2\right\}\right].$$

Here $\sum_{k=1}^p (X_{1k} - X_{2k})^2 | \boldsymbol{X}_1$ follows a non-central chi-square distribution with the non-central parameter $\sum_{k=1}^p X_{1k}^2$. We now show that this model satisfies Condition 9. Recall that

$$\begin{aligned} w'_{12} =& w_{12} - \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12}\right\} \\ & - \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_2}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{12}\right\} \\ & + \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top w_{12}\right\} \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) \\ =: & W_1 - W_2 - W_3 + W_4. \end{aligned}$$

Since ε_1 and $\boldsymbol{X}_1$ are independent, we just prove $E\left(w'^4_{12}\right) = O(1)$. And we will show

$$E\left[\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)\right\}^4\right] = O(1)$$

in the process of our proof.

By Cauchy-Schwarz inequality, we need to show w_1, W_2, W_3, *and* W_4 *have finite fourth-order moments. Since* $W_1 = w_{12} \leq 1$, $E(w_{12}^4) \leq 1$. *Note that*

$$\begin{aligned}
W_2 =& \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_3) w_{13} \right\} \\
=& \sum_{l=1}^{p} X_{2l} E_{\boldsymbol{X}_1} \left[X_{3l} \exp \left\{ -\frac{1}{p} \sum_{k=1}^{p} (X_{1k} - X_{3k})^2 \right\} \right] \\
=& \sum_{l=1}^{p} X_{2l} \left(\frac{p}{p+2} \right)^{p/2} \frac{2X_{1l}}{p+2} \exp \left(-\frac{\sum_{k=1}^{p} X_{1k}^2}{p+2} \right) \\
=& \frac{1}{p} \sum_{l=1}^{p} 2X_{1l} X_{2l} \exp \left(-\frac{\sum_{k=1}^{p} X_{1k}^2}{p+2} \right) \left(\frac{p}{p+2} \right)^{p/2+1} \\
\leq& \frac{2}{p} \sum_{l=1}^{p} \left| X_{1l} X_{2l} \right|, \qquad \text{(S1.2)}
\end{aligned}$$

where the third line is from (S1.5) below. Since $E\left\{(X_{1l}X_{2l})^4\right\} = 9$, $E\left(W_2^4\right)$ *is smaller than a positive constant. Similarly,* $E\left(W_3^4\right)$ *can be bounded by a positive constant. For* W_4, *we have, combining the presentation of* W_2 *in (S1.2),*

$$\begin{aligned}
W_4 =& \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top w_{12} \right\} \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) \\
=& \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top E \left\{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top w_{12} \right\} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) \\
=& \sum_{h=1}^{p} \sum_{l=1}^{p} X_{1h} X_{2l} E \left[X_{3h} X_{4l} \exp \left\{ -\frac{1}{p} \sum_{k=1}^{p} (X_{3k} - X_{4k})^2 \right\} \right] \\
=& \sum_{h=1}^{p} \sum_{l=1}^{p} X_{1h} X_{2l} E \left(X_{3h} E_{\boldsymbol{X}_3} \left[X_{4l} \exp \left\{ -\frac{1}{p} \sum_{k=1}^{p} (X_{3k} - X_{4k})^2 \right\} \right] \right) \\
=& \sum_{h=1}^{p} \sum_{l=1}^{p} X_{1h} X_{2l} E \left\{ X_{3h} \left(\frac{p}{p+2} \right)^{p/2} \frac{2X_{3l}}{p+2} \exp \left(-\frac{\sum_{k=1}^{p} X_{3k}^2}{p+2} \right) \right\}
\end{aligned}$$

$$=\frac{2}{p}\sum_{h=1}^{p}\sum_{l=1}^{p}X_{1h}X_{2l}E\left\{X_{3h}X_{3l}\exp\left(-\frac{\sum_{k=1}^{p}X_{3k}^2}{p+2}\right)\right\}\left(\frac{p}{p+2}\right)^{p/2+1}.$$

Further, when $h \neq l$, the independence between the zero mean variables X_{3h} and X_{3l} yields $E\left\{X_{3h}X_{3l}\exp\left(-\frac{\sum_{k=1}^{p}X_{3k}^2}{p+2}\right)\right\}\left(\frac{p}{p+2}\right)^{p/2+1}=0$, hence

$$\begin{aligned}W_4 =&\frac{2}{p}\sum_{h=1}^{p}X_{1h}^2E\left\{X_{3h}^2\exp\left(-\frac{\sum_{k=1}^{p}X_{3k}^2}{p+2}\right)\right\}\left(\frac{p}{p+2}\right)^{p/2+1}\\ <&\frac{2}{p}\sum_{h=1}^{p}X_{1h}^2E\left(X_{3h}^2\right)=\frac{2}{p}\sum_{h=1}^{p}X_{1h}^2.\end{aligned}$$

Note that $E(X_{1h}^8)=105<\infty$. Hence $E(W_4^4)$ is bounded by a constant. Then Condition 9 is satisfied. To complete the above calculation, we give the details below about the required integral calculations.

Set $u=X_{11}$, $\lambda=\sum_{k=2}^{p}X_{1k}^2$. We have

$$\begin{aligned}&E_{\boldsymbol{X}_1}\left[X_{21}\exp\left\{-\frac{1}{p}(u-X_{21})^2\right\}\right]\\ =&\int_0^\infty x\exp\left\{-\frac{1}{p}\left(x-u\right)^2\right\}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)dx\\ =&\frac{1}{\sqrt{2\pi}}\int_0^\infty x\exp\left\{-\frac{(p+2)\left(x^2-\frac{4u}{p+2}x+\frac{2u^2}{p+2}\right)}{2p}\right\}dx\\ =&\frac{1}{\sqrt{2\pi}}\int_0^\infty x\exp\left\{-\frac{(p+2)\left(x-\frac{2u}{p+2}\right)^2}{2p}\right\}dx\exp\left\{\frac{2u^2}{p\left(p+2\right)}-\frac{u^2}{p}\right\}\\ =&\frac{\sqrt{p/\left(p+2\right)}}{\sqrt{2\pi p/\left(p+2\right)}}\int_0^\infty x\exp\left\{-\frac{\left(x-\frac{2u}{p+2}\right)^2}{2p/\left(p+2\right)}\right\}dx\exp\left\{\frac{2u^2}{p\left(p+2\right)}-\frac{u^2}{p}\right\}\end{aligned}$$

$$=\sqrt{\frac{p}{p+2}}\frac{2u}{p+2}\exp\left\{\frac{2u^2}{p\,(p+2)}-\frac{u^2}{p}\right\}=\sqrt{\frac{p}{p+2}}\frac{2u}{p+2}\exp\left(-\frac{u^2}{p+2}\right), \tag{S1.3}$$

and

$$\begin{aligned}
&E_{\boldsymbol{X}_1}\left[\exp\left\{-\frac{1}{p}\sum_{k=2}^{p}(X_{1k}-X_{2k})^2\right\}\right]\\
&=\int_0^\infty \exp\left(-\frac{x}{p}\right)\sum_{j=0}^{\infty}\exp\left(-\frac{\lambda}{2}\right)\frac{(\lambda/2)^j}{j!}\\
&\quad\times\frac{1}{2^{j+(p-1)/2}\Gamma\left(j+(p-1)/2\right)}x^{j+(p-1)/2-1}\exp\left(-\frac{1}{2}x\right)dx\\
&=\exp\left(-\frac{\lambda}{2}\right)\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!}\int_0^\infty\frac{1}{2^{j+(p-1)/2}\Gamma\left(j+(p-1)/2\right)}x^{j+(p-1)/2-1}\\
&\quad\times\exp\left(-\frac{p+2}{2p}x\right)dx\\
&=\exp\left(-\frac{\lambda}{2}\right)\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!\left\{(p+2)/p\right\}^{j+(p-1)/2}}\\
&\quad\times\int_0^\infty\frac{\left\{(p+2)/p\right\}^{j+(p-1)/2}}{2^{j+(p-1)/2}\Gamma\left(j+(p-1)/2\right)}x^{j+(p-1)/2-1}\exp\left(-\frac{p+2}{2p}x\right)dx\\
&=\exp\left(-\frac{\lambda}{2}\right)\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!\left\{(p+2)/p\right\}^{j+(p-1)/2}}=\exp\left(-\frac{\lambda}{p+2}\right)\frac{1}{\left\{(p+2)/p\right\}^{(p-1)/2}}\\
&=\exp\left(-\frac{\sum_{k=2}^{p}X_{1k}^2}{p+2}\right)\left\{\frac{1}{(1+2/p)^{p/2-1/2}}\right\}.
\end{aligned} \tag{S1.4}$$

From (S1.3) and (S1.4), we have

$$\begin{aligned}
&\sum_{l=1}^{p}X_{2l}E_{\boldsymbol{X}_1}\left[X_{3l}\exp\left\{-\frac{1}{p}\sum_{k=1}^{p}(X_{1k}-X_{3k})^2\right\}\right]\\
&=\sum_{l=1}^{p}X_{2l}E_{\boldsymbol{X}_1}\left[X_{3l}\exp\left\{-\frac{1}{p}(X_{1l}-X_{3l})^2\right\}\right]E_{\boldsymbol{X}_1}\left[\exp\left\{-\frac{1}{p}\sum_{k=1,k\neq l}^{p}(X_{1k}-X_{3k})^2\right\}\right]
\end{aligned}$$

$$= \sum_{l=1}^{p} X_{2l} \sqrt{\frac{p}{p+2}} \frac{2X_{1l}}{p+2} \exp\left(-\frac{X_{1l}^2}{p+2}\right) \exp\left(-\frac{\sum_{k=1,k\neq l}^{p} X_{1k}^2}{p+2}\right) \left[\frac{1}{\{(p+2)/p\}^{p/2-1/2}}\right]$$

$$= \sum_{l=1}^{p} X_{2l} \left(\frac{p}{p+2}\right)^{p/2} \frac{2X_{1l}}{p+2} \exp\left(-\frac{\sum_{k=1}^{p} X_{1k}^2}{p+2}\right). \tag{S1.5}$$

Recall the definition of $\boldsymbol{\Sigma}_l = E\left\{l^2(\boldsymbol{X})\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}\right)\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}\right)^\top\right\}$ before Condition 7.10 in the main context.??? Under the global hypothesis, $l(\boldsymbol{X}) = m(\boldsymbol{X}) - g(\boldsymbol{\theta}^*, \boldsymbol{X})$.

Condition 10. *$E\{l^4(\boldsymbol{X})\} = O(1)$. Further, there are two constants a_l, b_l not depending on n and p such that*

$$0 < a_l \leq \lambda_{\min}(\boldsymbol{\Sigma}_l) \leq \lambda_{\min}(\boldsymbol{\Sigma}_l) \leq b_l < \infty.$$

Recall the sequence of local alternatives,

$$H_{1n} : Y = g(\boldsymbol{\theta}_0, \boldsymbol{X}) + \delta_n l(\boldsymbol{X}) + \varepsilon$$

and

$$\begin{aligned} H(\boldsymbol{X}_1) =& E_{\boldsymbol{X}_1}\{l(\boldsymbol{X}_2)w_{12}\} - E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)w_{12}\}^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\} \\ & - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\} \\ & + \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\}. \end{aligned}$$

in Subsection 4.2 of the main context. We repeat it here for convenience.

Condition 11. *$E\{|\varepsilon_1 H(\boldsymbol{X}_1)|^2\} \geq C_5 > 0$ and $E\{|\varepsilon_1 H(\boldsymbol{X}_1)|^4\} \leq C_6 < \infty$, where C_5 and C_6 are two constants unrelated to n and p. $H(\boldsymbol{X}_1) \neq 0, a.s.$*

S2 Proofs of the theorems

S2.1 A brief summary of the main results

To establish the asymptotic distributions under both the null and alternative hypotheses, we first employ a careful decomposition that isolates a leading term involving the errors and a remainder term that captures the estimation effect. The crux of the proof lies in showing that this remainder term is asymptotically negligible. To this end, a case-by-case analysis of these remainder terms is required to determine when they become negligible asymptotically. We address these remainder terms through two principal strategies. First, we apply a Taylor expansion to approximate them by sums of i.i.d. variables, then establish their probabilistic bounds and convergence rates using Markov's or the Cauchy–Schwarz inequality. Second, we control the remainder directly by exploiting the distance between the parameter and its estimator. Leveraging the estimator's convergence rate and its asymptotic linear representation, we rigorously prove the asymptotic negligibility of these terms. **To enhance readability, we provide a brief overview of the objective of each lemma and theorem presented in this section.**

The test statistic is the ratio of two random sequences and involves the

estimator $\hat{\boldsymbol{\theta}}$ of the parameter $\boldsymbol{\theta}^*$ ($\boldsymbol{\theta}_0$). The randomness of these variables causes the complication of the proofs of the main theorems that are rather lengthy. Thus, we separate the details of the proofs into several lemmas in summary of their purposes in the proofs before going to the details. First, the deviation between the parameter estimator and its true value causes the difference between the residual and the error term.

To analyse the deviation of the parameter estimator, Lemmas 1 and 2 provide the uniformly asymptotically linear representation of $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*$ such that the analysis for the difference between the residual and the error term can be implemented. To derive the asymptotic properties of the proposed test statistic, we decompose its numerator and denominator into several parts, respectively. In the decomposition, we note that two terms are very important for our analysis, so we give the limiting properties of these two terms in Lemmas 3 and 4. Lemma 5 gives the limiting distribution of the numerator under the null hypothesis. Besides, Lemmas 5 and 6 are used to prove Equation (4) in the main context, which is a primary motivation for constructing the statistic. Using the martingale central limit theorem, Lemmas 7 and 8 show the asymptotic normality of a martingale sequence, which is crucial for proving that the denominator in the test statistic has a positive lower bound in probability. These two lemmas are useful in the

proofs of Lemmas 9, 10 and 13. Lemma 9 is used to prove the asymptotic normality, under the null hypothesis, of V_n defined in the main context. Concretely speaking, incorporating Lemma 9 with Lemma 13, we can obtain that the denominator in the statistic can be used as an estimator of the conditional expectation of the numerator in the statistic. See Corollary 3 for details. Note that the numerator and denominator are random, and we need to deal with the conditional distribution of the test statistic when the second part of the sample is given. To this end, the Berry-Essen inequality is in need. Lemmas 10 and 13 are to check the condition the inequality requires. Under the alternative hypothesis Lemmas 11 and 12 give the limiting distribution of the numerator. Lemmas 14 and 15 are useful in the proofs for Lemmas 16 and 18. Further, Lemmas 16, 17, and 18 show the limiting properties of the denominator. Lemma 19 restates the classical Berry-Essen inequality. Finally, we can easily finish the proofs for **Theorem 1** and **Theorem 2** in the main context, which correspond to the limiting properties of the statistic under the null and the alternative hypotheses, respectively. To more easily access the details of the proofs, we present some commonly used notations and definitions again here, some of which have been defined in the paper.

First, $\|\cdot\|$ denotes as the L_2 norm and $E_{\boldsymbol{X}}(\cdot)$ as the conditional expec-

tation given $\boldsymbol{X}$, such as $E_{\boldsymbol{X}_1}(\cdot) = E(\cdot|\boldsymbol{X}_1)$ and $E_{\mathcal{N}_2}(\cdot) = E(\cdot|\mathcal{N}_2)$. Further, let $\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top$, where

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \sum_{i=1}^{n} \{Y_i - g(\boldsymbol{\theta}, \boldsymbol{X}_i)\}^2 .$$

Then the weight function $\tilde{w}_i(\mathcal{N}_2)$ and its estimator are written as

$$\tilde{w}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[w_{ij} - \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \left\{ \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_1\right) w_{1j} \right\} \right]$$

and

$$\tilde{\hat{w}}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[w_{ij} - \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right],$$

Besides, for the sake of convenience, define

$$\tilde{w}_i^{(0)}(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[w_{ij} - \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}_*^{-1} E_{\boldsymbol{X}_j} \{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{1j} \} \right]$$

in the proofs for the results under the global alternative hypothesis, where $\boldsymbol{\Sigma}_* = E\left\{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top \right\}$.

Recall the null hypothesis,

$$H_0 : \mathbb{P}\{E(Y|\boldsymbol{X}) = g(\boldsymbol{\theta}_0, \boldsymbol{X})\} = 1 \text{ for some } \boldsymbol{\theta}_0 \in \boldsymbol{\Theta}, \tag{S2.6}$$

and the alternative hypothesis,

$$H_1 : \mathbb{P}\{E(Y|\boldsymbol{X}) = g(\boldsymbol{\theta}, \boldsymbol{X})\} < 1 \text{ for all } \boldsymbol{\theta} \in \boldsymbol{\Theta}.$$

H_1 can be rewritten as

$$H_{1n} : Y = g(\boldsymbol{\theta}_0, \boldsymbol{X}) + \delta_n l(\boldsymbol{X}) + \varepsilon. \tag{S2.7}$$

When $\delta_n \to 0$ as $n \to \infty$, (S2.7) corresponds to the local alternative hypothesis; and $\delta_n \equiv c > 0$ to the global alternative hypothesis.

The following Lemmas 1 and 2 are derived form Theorem 1 and Theorem 2 in Tan and Zhu (2022).

Lemma 1. *Suppose that Conditions 1-4 hold. If* $p^2/n \to 0$, *then it follows that* $\hat{\boldsymbol{\theta}}$ *is a norm consistent estimator of* $\boldsymbol{\theta}^*$ *in the sense that* $\left\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\right\| = O_p\left(\sqrt{p/n}\right)$. *Furthermore, if* $(p^3 \log n)/n \to 0$, *then additionally, assuming that Condition 5 holds, we have*

$$n^{1/2}\boldsymbol{\alpha}^\top\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*\right) = n^{-1/2}\sum_{i=1}^{n}\left\{Y_i - g(\boldsymbol{\theta}^*, \boldsymbol{X}_i)\right\}\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i) + o_p(1), \tag{S2.8}$$

where the remaining term $o_p(1)$ *is uniformly over* $\boldsymbol{\alpha} \in \mathcal{S}^{p-1}$ *which is the unit sphere in* $\mathbb{R}^p$.

Lemma 2. *Suppose that Conditions 1-5 and 11 hold. Under* H_{1n}, *if* $\delta_n = 1/\sqrt{n}$ *or* $\delta_n\sqrt{n} \to \infty$, *then when* $p\delta_n \to 0$, $\hat{\boldsymbol{\theta}}$ *is a norm consistent estimator for* $\boldsymbol{\theta}_0$ *with* $\left\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right\| = O_p\left(\sqrt{p}\delta_n\right)$. *If* $\delta_n\sqrt{n} \to 0$ *and* $p^2/n \to 0$, $\hat{\boldsymbol{\theta}}$ *is a norm consistent estimator for* $\boldsymbol{\theta}_0$ *with* $\left\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right\| = O_p\left(\sqrt{p/n}\right)$. *Moreover, when* $p^3 \log n/n \to 0$ *and* $np^3 \log n\delta_n^4 \to 0$, *we have*

$$\begin{aligned}\sqrt{n}\boldsymbol{\alpha}^\top\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) =& \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\varepsilon_j\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \\ &+ \sqrt{n}\delta_n\big(\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\} + o_p(1)\big),\end{aligned} \tag{S2.9}$$

where the term $o_p(1)$ *is uniform over* $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{S}}^{p-1}$.

By taking advantage of the asymptotic linear representation of $\hat{\boldsymbol{\theta}}$ and the Taylor expansion, we can derive the limiting properties of the term $\sum_{j=n_1+1}^{n} \left\{ g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) - g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \right\} w_{ij}$ for any fixed $1 \leq i \leq n_1$, which is significant of both the numerator and the denominator in our statistic. To avoid repeatedly invoking the asymptotic representation and Taylor expansion whenever discussing this term in subsequent discussions, we summarize the following lemma for ease of future reference. Without loss of generality, consider $i = 1$ in the following statement.

Lemma 3. *Under Conditions 1-7, if* $p^3 \log n_2 / n_2 \to 0$, *and* $\hat{\boldsymbol{\theta}}_2$ *is estimated by* $(\boldsymbol{X}_j, Y_j), j = n_1 + 1, \cdots, n$, *then we have following results.*

(a) Under the null hypothesis H_{1n} *in (S2.7) with* $\delta_n = 0$,

$$\begin{aligned} &\sum_{j=n_1+1}^{n} \left\{ g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) - g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \right\} w_{1j} \\ =& \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1} \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12} \} + \sqrt{n_2} R_1, \end{aligned}$$

where $R_1 = o_p(1)$ *and* $E_{\mathcal{N}_2}(R_1^4) = o_p(1)$.

(b) Under the global alternative hypothesis H_{1n} *in (S2.7) with* $\delta_n = 1$,

$$\sum_{j=n_1+1}^{n} \left\{ g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) - g(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\} w_{1j}$$

$$= \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12}\}$$
$$+ \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12}\} + \sqrt{n_2}\mathcal{S}_1,$$

where $l(\boldsymbol{X}_j) = m(\boldsymbol{X}_j) - g(\boldsymbol{\theta}^*, \boldsymbol{X}_j)$, $\mathcal{S}_1 = o_p(1)$ *and* $E_{\mathcal{N}_2}(\mathcal{S}_1^4) = o_p(1)$.

(c) Under the local alternative hypothesis H_{1n} *in (S2.7), we have*

$$\sum_{j=n_1+1}^{n} \left\{ g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) - g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \right\} w_{1j}$$
$$= \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12}\}$$
$$+ \sum_{j=n_1+1}^{n} \delta_n \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\} w_{1j} + \sqrt{n_2}\mathcal{R}_1,$$

where $\mathcal{R}_1 = o_p(1)$ *and* $E_{\mathcal{N}_2}(\mathcal{R}_1^4) = o_p(1)$.

Proof of Lemma 3. **Prove cases (a) and (b).** From Lemma 1, we derive that under the null hypothesis and the global alternative hypothesis,

$$\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0 = \frac{1}{n_2} \sum_{j=n_1+1}^{n} \{Y_j - g(\boldsymbol{\theta}^*, \boldsymbol{X}_j)\} \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + \boldsymbol{R}_{2n}, \qquad \text{(S2.10)}$$

where $\|\boldsymbol{R}_{2n}\| = o_p\left(1/\sqrt{n_2}\right)$. According to the asymptotically linear representation of $\boldsymbol{\theta}$ (equation (S2.8) in Lemma 1), we have, by Taylor expansion, there is a $\tilde{\boldsymbol{\theta}}^*$ between $\hat{\boldsymbol{\theta}}_2$ and $\tilde{\boldsymbol{\theta}}^*$,

$$\sum_{j=n_1+1}^{n} \left\{ g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) - g(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\} w_{1j}$$
$$= \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \sum_{j=n_1+1}^{n} \dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) w_{1j}$$

$$
\begin{aligned}
&= \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j} \\
&\quad + \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \sum_{j=n_1+1}^{n} \left\{ \dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) w_{1j} - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) w_{1j} \right\} \\
&= \frac{1}{n_2} \sum_{j_1=n_1+1}^{n} \left\{ Y_{j_1} - g(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_1}) \right\} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_1})^\top \boldsymbol{\Sigma}^{-1} \sum_{j_2=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_2}) w_{1j_2} \\
&\quad + \boldsymbol{R}_{2n}^\top \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j} \\
&\quad + \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \sum_{j=n_1+1}^{n} \left\{ \dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) w_{1j} - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) w_{1j} \right\} \\
&= \sum_{j=n_1+1}^{n} \left\{ Y_j - g(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1} \left\{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12} \right\} \\
&\quad + \sum_{j=n_1+1}^{n} \left\{ Y_j - g(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{S}_{1n_2} \\
&\quad + \boldsymbol{R}_{2n}^\top \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j} \\
&\quad + \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \sum_{j=n_1+1}^{n} \left\{ \dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) w_{1j} - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) w_{1j} \right\}, \\
&=: \sum_{j=n_1+1}^{n} \left\{ Y_j - g(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_1} \left\{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12} \right\} + \sqrt{n_2} R_1,
\end{aligned}
\tag{S2.11}
$$

where $\boldsymbol{S}_{1n_2} = \frac{1}{n_2} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j} - E_{\boldsymbol{X}_1} \left\{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2) w_{12} \right\}$. Under the null hypothesis, $Y_1 - g(\boldsymbol{\theta}^*, \boldsymbol{X}_1) = \varepsilon_1$, then in case (a),

$$
\begin{aligned}
R_1 =& \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{S}_{1n_2} + \boldsymbol{R}_{2n}^\top \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j} \\
&+ \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ \dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) w_{1j} - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) w_{1j} \right\}
\end{aligned}
$$

$$=: R'_1 + R''_1 + R'''_1.$$

Under the global alternative hypothesis with $\delta_n = 1$ and $Y_1 - g(\boldsymbol{\theta}^*, \boldsymbol{X}_1) = \varepsilon_1 + l(\boldsymbol{X}_1)$, in case (b),

$$\begin{aligned}
\mathcal{S}_1 =& \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{S}_{1n_2} \\
&+ \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{S}_{1n_2} \\
&+ \boldsymbol{R}_{2n}^\top \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j} \\
&+ \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ \dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) w_{1j} - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) w_{1j} \right\} \\
=: & R'_1 + \mathcal{S}'_1 + R''_1 + R'''_1.
\end{aligned}$$

Recall $E_{\mathcal{N}_2}(\cdot) = E(\cdot|\mathcal{N}_2)$ means given $\mathcal{N}_2$, the conditional expectation over other random variable(s). For case (a), we first show $E_{\mathcal{N}_2}\left(R_1^4\right) = o_p(1)$ and use it to derive $R_1 = o_p(1)$. The reason is as follows: if $E_{\mathcal{N}_2}\left(R_1^4\right) = o_p(1)$, using Markov inequality, we know that for any $\eta > 0$, $\mathbb{P}\left(R_1^4 > \eta | \mathcal{N}_2\right) \leq E_{\mathcal{N}_2}\left(R_1^4\right)/\eta = o_p(1)$. Note that $\mathbb{P}\left(R_1^4 > \eta | \mathcal{N}_2\right) \leq 1$ is uniformly integrable. Thus, we conclude that $\mathbb{P}\left(R_1^4 > \eta\right) = E\left\{\mathbb{P}\left(R_1^4 > \eta | \mathcal{N}_2\right)\right\} = o(1)$. That is to say, $R_1^4 = o_p(1)$. For case (b), we use a similar strategy. To prove $E_{\mathcal{N}_2}\left(R_1^4\right) = o_p(1)$ and $E_{\mathcal{N}_2}\left(\mathcal{S}_1^4\right) = o_p(1)$, we show that $E_{\mathcal{N}_2}\left(R'^4_1\right) = o_p(1)$, $E_{\mathcal{N}_2}\left(R''^4_1\right) = o_p(1)$, $E_{\mathcal{N}_2}\left(R'''^4_1\right) = o_p(1)$ and $E_{\mathcal{N}_2}\left(\mathcal{S}'^4_1\right) = o_p(1)$ in the

following.

First, we prove $E_{\mathcal{N}_2}\left(R_1'^4\right) = o_p(1)$. Note that

$$\begin{aligned} E_{\mathcal{N}_2}\left(R_1'^4\right) =& E_{\mathcal{N}_2}\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{S}_{1n_2} \right\|^4 \\ \leq & E_{\mathcal{N}_2}\left\{ \left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \right\|^4 \left\|\boldsymbol{\Sigma}^{-1}\right\|^4 \left\|\boldsymbol{S}_{1n_2}\right\|^4 \right\} \\ = & \left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \right\|^4 \left\|\boldsymbol{\Sigma}^{-1}\right\|^4 E_{\mathcal{N}_2}\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right). \quad \text{(S2.12)} \end{aligned}$$

We now show that

$$E_{\mathcal{N}_2}\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right) = O_p\left(p^2/n_2^2\right)$$

and

$$\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \right\|^4 = O_p\left(p^2\right).$$

For $E_{\mathcal{N}_2}\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right)$, according to Markov inequality, if we can show $E\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right) = O\left(p^2/n_2^2\right)$, then $E_{\mathcal{N}_2}\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right) = O_p\left(p^2/n_2^2\right)$ is a direct result. Recall that $E_{\boldsymbol{X}}(\cdot) = E(\cdot|\boldsymbol{X})$. Under Condition 2, using Cauchy-Schwarz inequality,

$$\begin{aligned} E_{\boldsymbol{X}_1}\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right) =& E_{\boldsymbol{X}_1}\left(\sum_{k=1}^{p} \left[\frac{1}{n_2} \sum_{j=n_1+1}^{n} \dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{1j} - E_{\boldsymbol{X}_1}\left\{ \dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12} \right\} \right]^2 \right)^2 \\ =& \sum_{k=1}^{p} \sum_{l=1}^{p} E_{\boldsymbol{X}_1}\left[\frac{1}{n_2} \sum_{j=n_1+1}^{n} \dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{1j} - E_{\boldsymbol{X}_1}\left\{ \dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12} \right\} \right]^2 \\ & \times E_{\boldsymbol{X}_1}\left[\frac{1}{n_2} \sum_{j=n_1+1}^{n} \dot{g}_l(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{1j} - E_{\boldsymbol{X}_1}\left\{ \dot{g}_l(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12} \right\} \right]^2 \end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{k=1}^{p}\sum_{l=1}^{p}\frac{1}{n_2^2}E_{\boldsymbol{X}_1}\left\{\dot{g}_k^2(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}^2\right\}E_{\boldsymbol{X}_1}\left\{\dot{g}_l^2(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}^2\right\}\\
&\leq \sum_{k=1}^{p}\sum_{l=1}^{p}\frac{1}{n_2^2}E^{1/2}\left\{\dot{g}_k^4(\boldsymbol{\theta}_0,\boldsymbol{X}_2)\right\}E^{1/2}\left\{\dot{g}_l^4(\boldsymbol{\theta}_0,\boldsymbol{X}_2)\right\}E_{\boldsymbol{X}_1}\left(w_{12}^4\right)\\
&\leq \frac{p^2}{n_2^2}E\left\{F^4(\boldsymbol{X}_2)\right\}E_{\boldsymbol{X}_1}\left(w_{12}^4\right),
\end{aligned}
$$

where the last inequality is derived from the facts $|\dot{g}_k(\boldsymbol{\theta}_0,\boldsymbol{X}_2)|\leq F(\boldsymbol{X}_2)$ for any k. Under Condition 7, $E\left(w_{12}^4\right)\leq C_1$. Thus, integrating $E_{\boldsymbol{X}_1}\left(\|\boldsymbol{S}_{1n_2}\|^4\right)$ over $\boldsymbol{X}_1$, we have

$$
E\left(\|\boldsymbol{S}_{1n_2}\|^4\right)\leq\frac{p^2}{n_2^2}E\left\{F^4(\boldsymbol{X}_2)\right\}E\left(w_{12}^4\right)=O\left(p^2/n_2^2\right). \tag{S2.13}
$$

Therefore, $E_{\mathcal{N}_2}\left(\|\boldsymbol{S}_{1n_2}\|^4\right)=O_p\left(p^2/n_2^2\right)$. For $\left\|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^4$, since

$$
\begin{aligned}
&E\left\|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^2\\
=&E\left\{\frac{1}{n_2}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n}\varepsilon_{j_1}\varepsilon_{j_2}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_1})^\top\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_2})\right\}\\
=&E\left\{\frac{1}{n_2}\sum_{j=n_1+1}^{n}\varepsilon_j^2\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^2\right\}=O\left(p\right),
\end{aligned}
$$

we can obtain that $\left\|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^2=O_p\left(p\right)$ by applying Markov inequality. Thus, $\left\|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^4=O_p\left(p^2\right)$.

Under Condition 4, $\left\|\boldsymbol{\Sigma}^{-1}\right\|=O\left(1\right)$. Then according to (S2.12), as $p^3/n_2\to 0$, we have

$$
E_{\mathcal{N}_2}\left(R_1'^4\right)=O_p\left(p^2\right)O\left(1\right)O_p\left(p^2/n_2^2\right)=o_p\left(1\right).
$$

Second, prove $E_{\mathcal{N}_2}\left(\mathcal{S}_1'^4\right) = o_p(1)$. Similar to (S2.12), we have

$$\begin{aligned} E_{\mathcal{N}_2}\left(\mathcal{S}_1'^4\right) =& E_{\mathcal{N}_2}\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{S}_{1n_2} \right\|^4 \\ \leq& E_{\mathcal{N}_2}\left\{ \left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\|^4 \left\|\boldsymbol{\Sigma}^{-1}\right\|^4 \left\|\boldsymbol{S}_{1n_2}\right\|^4 \right\} \\ =& \left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\|^4 \left\|\boldsymbol{\Sigma}^{-1}\right\|^4 E_{\mathcal{N}_2}\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right). \end{aligned} \tag{S2.14}$$

We have already shown $E_{\mathcal{N}_2}\left(\left\|\boldsymbol{S}_{1n_2}\right\|^4\right) = O_p(p^2/n_2^2)$. We then prove $\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\|^4 = O_p(p^2)$ in the following. Here, we use a crucial property: under the global alternative hypothesis in (S2.7) with fixed $\delta_n = 1$, $E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\right\} = \mathbf{0}$. This is because

$$\begin{aligned} \left.\frac{\partial E\left\{Y_1 - g\left(\boldsymbol{\theta}, \boldsymbol{X}_1\right)\right\}^2}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} =& \left.\frac{\partial E\left\{m(\boldsymbol{X}_1) - g\left(\boldsymbol{\theta}, \boldsymbol{X}_1\right)\right\}^2}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \\ =& -2E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\right\} = \mathbf{0}. \end{aligned} \tag{S2.15}$$

Thus, for any $j_1 \neq j_2$, we have $E\left\{l(\boldsymbol{X}_{j_1})l(\boldsymbol{X}_{j_2})\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_1})^\top \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_2})\right\} = 0$. Then it follows that

$$\begin{aligned} & E\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\|^2 \\ =& E\left\{ \frac{1}{n_2} \sum_{j_1=n_1+1}^{n} \sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_1})l(\boldsymbol{X}_{j_2})\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_1})^\top \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_2}) \right\} \\ =& E\left\{ \frac{1}{n_2} \sum_{j=n_1+1}^{n} l^2(\boldsymbol{X}_j) \left\|\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)\right\|^2 \right\} = O(p), \end{aligned}$$

this yields that $\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \right\|^4 = O_p\left(p^2\right)$. Then according to (S2.14), we derive that

$$E_{\mathcal{N}_2}\left(\mathcal{S}_1'^4\right) = O_p\left(p^2\right) O\left(1\right) O_p\left(p^2/n_2^2\right) = o_p\left(1\right).$$

Third, we try to prove $E_{\mathcal{N}_2}\left(R_1''^4\right) = o_p\left(1\right)$. Recall $\|\boldsymbol{R}_{2n}\| = o_p\left(1/\sqrt{n_2}\right)$ in (S2.10) and $\boldsymbol{R}_{2n}$ is $\mathcal{N}_2$ measurable. From Conditions 4 and 7, we know $E\left[\|E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)w_{12}\right\}\|^4\right] \leq C_2$. See Remark 1 for details. As $E_{\mathcal{N}_2}\left(\|\boldsymbol{S}_{1n_2}\|^4\right) = O_p\left(p^2/n_2^2\right)$, we have

$$\begin{aligned}
E_{\mathcal{N}_2}\left(R_1''^4\right) =& E_{\mathcal{N}_2}\left[\left\{\boldsymbol{R}_{2n}^\top \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)w_{1j}\right\}^4\right] \\
\leq& \|\boldsymbol{R}_{2n}\|^4 E_{\mathcal{N}_2}\left\{\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)w_{1j}\right\|^4\right\} \\
\leq& 16n_2^2 \|\boldsymbol{R}_{2n}\|^4 E\left[\|E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)w_{12}\right\}\|^4\right] \\
&+ 16n_2^2 \|\boldsymbol{R}_{2n}\|^4 E_{\mathcal{N}_2}\left(\|\boldsymbol{S}_{1n_2}\|^4\right) \\
=& n_2^2 o_p\left(1/n_2^2\right) O\left(1\right) + n_2^2 o_p\left(1/n_2^2\right) O_p\left(p^2/n_2^2\right) = o_p\left(1\right), \quad \text{(S2.16)}
\end{aligned}$$

where $\boldsymbol{S}_{1n_2} = \frac{1}{n_2}\sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)w_{1j} - E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)w_{12}\right\}$.

Fourth, we start to discuss $E_{\mathcal{N}_2}\left(R_{11}'''^4\right) = o_p\left(1\right)$. According to Condition 6, as $\tilde{\boldsymbol{\theta}}^*$ lies between $\boldsymbol{\theta}^*$ and $\hat{\boldsymbol{\theta}}_2$, we have

$$\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right)\right\| \leq \sqrt{p}\left\|\tilde{\boldsymbol{\theta}}^* - \boldsymbol{\theta}^*\right\| L(\boldsymbol{X}_j) \leq \sqrt{p}\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right\| L(\boldsymbol{X}_j)$$

and $E\{L^4(\boldsymbol{X}_1)\} = O(1)$. Recall $\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right\| = O_p\left(\sqrt{p/n_2}\right)$. Thus,

$$
\begin{aligned}
& E_{\mathcal{N}_2}\left(R'''^4_1\right) \\
=& E_{\mathcal{N}_2}\left(\left[\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right)^\top \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) w_{1j} - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) w_{1j}\right\}\right]^4\right) \\
\leq & \left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right\|^4 E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right)\right\| w_{1j}\right\}^4\right] \\
\leq & \left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right\|^4 \left\{\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right)\right\|\right\}^4 \\
\leq & \left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right\|^4 \left\{\sqrt{p}\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} L(\boldsymbol{X}_j)\right\}^4 \\
=& p^2\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}^*\right\|^8 \left\{\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} L(\boldsymbol{X}_j)\right\}^4 = p^2 O_p\left(p^4/n_2^4\right) O_p\left(n_2^2\right) = o_p(1).
\end{aligned}
\tag{S2.17}
$$

Combining the results above, we have $E_{\mathcal{N}_2}(R_1^4) = o_p(1)$ and $E_{\mathcal{N}_2}(\mathcal{S}_1^4) = o_p(1)$. By Markov inequality, we know $R_1 = o_p(1)$ and $\mathcal{S}_1 = o_p(1)$.

The proofs for case (a) and case (b) is completed.

Prove case (c). Using the asymptotically linear representation of $\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0$ in Lemma 2, we derive that under the local alternative hypothesis in (S2.7),

$$
\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0 = \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + \delta_n \left[\boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} + \boldsymbol{\gamma}_{2n}\right],
$$

where $\boldsymbol{\gamma}_{2n} = o_p(1)$, and $\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\| = O_p\left(\max\left\{\sqrt{p/n_2}, \sqrt{p}\delta_n\right\}\right)$. By Tay-

lor expansion of $g(\cdot,\cdot)$ at $\boldsymbol{\theta}=\boldsymbol{\theta}_0$, there is a $\tilde{\boldsymbol{\theta}}_0$ between $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}_2$,

$$
\begin{aligned}
&\sum_{j=n_1+1}^{n}\left\{g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)-g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}w_{1j}\\
=&\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)w_{1j}\\
&+\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\sum_{j=n_1+1}^{n}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_0,\boldsymbol{X}_j\right)w_{1j}-\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)w_{1j}\right\}\\
=&\frac{1}{n_2}\sum_{j_1=n_1+1}^{n}\varepsilon_{j_1}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_1})^{\top}\boldsymbol{\Sigma}^{-1}\sum_{j_2=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_2})w_{1j_2}\\
&+\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{1j}\\
&+\delta_n\boldsymbol{\gamma}_{2n}^{\top}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)w_{1j}\\
&+\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\sum_{j=n_1+1}^{n}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_0,\boldsymbol{X}_j\right)w_{1j}-\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)w_{1j}\right\}\\
=&\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\\
&+\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{1j}\\
&+\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}_{1n_2}+\delta_n\boldsymbol{\gamma}_{2n}^{\top}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)w_{1j}\\
&+\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\sum_{j=n_1+1}^{n}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_0,\boldsymbol{X}_j\right)w_{1j}-\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)w_{1j}\right\},
\end{aligned}
\tag{S2.18}
$$

where $\boldsymbol{S}_{1n_2}=\frac{1}{n_2}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)w_{1j}-E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}$. Define

$$
\mathcal{R}_1=\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}_{1n_2}+\delta_n\boldsymbol{\gamma}_{2n}^{\top}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)w_{1j}
$$

$$+\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_0,\boldsymbol{X}_j\right)w_{1j}-\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)w_{1j}\right\}$$

$$=\mathcal{R}'_1+\mathcal{R}''_1+\mathcal{R}'''_1.$$

For case (c), we also prove that $E_{\mathcal{N}_2}\left(\mathcal{R}_1^4\right)=o_p\left(1\right)$, and then $\mathcal{R}_1=o_p\left(1\right)$ by Markov inequality. Similar to the proofs of case (a) and case (b), we have $E_{\mathcal{N}_2}\left(\mathcal{R}'^4_1\right)=o_p\left(1\right)$ and $E_{\mathcal{N}_2}\left(\mathcal{R}''^4_1\right)=o_p\left(1\right)$. Thus, we just need to prove $E_{\mathcal{N}_2}\left(\mathcal{R}'''^4_1\right)=o_p\left(1\right)$. Similar to (S2.17), we have

$$\begin{aligned}
&E_{\mathcal{N}_2}\left(\mathcal{R}'''^4_1\right)\\
=&E_{\mathcal{N}_2}\left(\left[\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_0,\boldsymbol{X}_j\right)w_{1j}-\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)w_{1j}\right\}\right]^4\right)\\
=&p^2\left\|\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right\|^8\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}L(\boldsymbol{X}_j)\right\}^4=o_p\left(1\right).
\end{aligned}$$

Hence, $E_{\mathcal{N}_2}\left(\mathcal{R}_1^4\right)=o_p\left(1\right)$ and then $\mathcal{R}_1=o_p\left(1\right)$. Further,

$$\begin{aligned}
&\sum_{j=n_1+1}^{n}\left\{g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)-g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}w_{1j}\\
=&\frac{1}{n_2}\sum_{j_1=n_1+1}^{n}\varepsilon_{j_1}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_1})^{\top}\boldsymbol{\Sigma}^{-1}\sum_{j_2=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_2})w_{1j_2}\\
&+\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\left[\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{1j}+\mathcal{R}_{11}\right].
\end{aligned}$$

The proof of case (c) is finished, and so is the whole proof of Lemma 3. □

Now, in the following lemma, we analyse another term, which is important for the asymptotic properties of the numerator in our statistic. Recall

the definition of $\boldsymbol{\Sigma}^*_{cov} = E\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top w_{12}\right\}$ in Condition 4.

Lemma 4. *Suppose that Conditions 1-8 hold,* $p^3 \log n_1/n_1 \to 0$ *and* $p^3 \log n_2/n_2 \to 0$. *Then the following conclusions hold:*

(a) Under the null hypothesis in (S2.6),

$$\begin{aligned}
&\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
=&\sum_{i=1}^{n_1}\varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\sum_{j=n_1+1}^{n}\varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + o_p\left(\sqrt{n_1 n_2}\right). \quad \text{(S2.19)}
\end{aligned}$$

(b) Under the local alternative hypothesis in (S2.7), with $p^3\delta_n^2 \log n \to 0$, *suppose that the additional conditions 10 and 11 hold. Then*

$$\begin{aligned}
&\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
=&\sum_{i=1}^{n_1}\varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\sum_{j=n_1+1}^{n}\varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
&+ n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n}\varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
&+ n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \\
&+ n_1 n_2 \delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
&+ o_p\left(\max\left\{\sqrt{n_1 n_2}, \sqrt{n_1} n_2 \delta_n, n_1\sqrt{n_2}\delta_n, n_1 n_2 \delta_n^2\right\}\right). \quad \text{(S2.20)}
\end{aligned}$$

Proof of Lemma 4. We need to prove case (b) as case (a) can be easily derived from similar arguments with $\delta_n = 0$ in (S2.7). Using Taylor

expansion, we decompose the above term into two parts:

$$\begin{aligned}
&\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\\
&=\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^{\top}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}\left(\tilde{\boldsymbol{\theta}}_1^{*},\boldsymbol{X}_i\right)\left\{g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)-g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}w_{ij}\\
&=\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^{\top}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\left\{g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)-g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}w_{ij}\\
&\quad+\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^{\top}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_1^{*},\boldsymbol{X}_i\right)-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\}\\
&\qquad\times\left\{g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)-g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}w_{ij}\\
&=:\Gamma_1+\Gamma_2,
\end{aligned}\tag{S2.21}$$

where $\tilde{\boldsymbol{\theta}}_1^{*}$ lies between $\hat{\boldsymbol{\theta}}_1$ and $\boldsymbol{\theta}_0$.

We can further decompose Γ_1 as

$$\begin{aligned}
\Gamma_1&=\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^{\top}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\left\{g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)-g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}w_{ij}\\
&=\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^{\top}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^{*},\boldsymbol{X}_j\right)^{\top}w_{ij}\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)\\
&=\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^{\top}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}w_{ij}\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)\\
&\quad+\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^{\top}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^{*},\boldsymbol{X}_j\right)-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}^{\top}w_{ij}\\
&\qquad\times\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)=:\Gamma_{11}+\Gamma_{12},
\end{aligned}\tag{S2.22}$$

where $\tilde{\boldsymbol{\theta}}_2^{*}$ lies between $\hat{\boldsymbol{\theta}}_2$ and $\boldsymbol{\theta}_0$.

Next, we analyse Γ_{11}, Γ_{12} and Γ_2. Divide the following proof into three

steps.

Step 1. Now we begin to deal with Γ_{11}. Γ_{11} can be decomposed as follows:

$$\begin{aligned}\Gamma_{11} =& \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\ &+ n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\ &+ n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \\ &+ n_1 n_2 \delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\ &+ o_p\left(\max\left\{\sqrt{n_1 n_2}, \sqrt{n_1} n_2 \delta_n, n_1 \sqrt{n_2} \delta_n, n_1 n_2 \delta_n^2\right\}\right). \end{aligned} \tag{S2.23}$$

To analyse Γ_{11} in (S2.22), we need to use the asymptotically linear representation of $\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0$ derived by Lemma 2,

$$\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0 = \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) + \delta_n \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} + \delta_n \boldsymbol{\gamma}_{1n} \tag{S2.24}$$

and

$$\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0 = \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + \delta_n \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} + \delta_n \boldsymbol{\gamma}_{2n}, \tag{S2.25}$$

where $\|\boldsymbol{\gamma}_{1n}\| = o_p\left(1/\sqrt{n_1}\right)$ and $\|\boldsymbol{\gamma}_{2n}\| = o_p\left(1/\sqrt{n_2}\right)$. Then according to

(S2.24) and (S2.25), we can write Γ_{11} as follows:

$$
\begin{aligned}
\Gamma_{11} = & \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
= & \frac{1}{n_1} \sum_{i_1=1}^{n_1} \varepsilon_{i_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_1})^\top \boldsymbol{\Sigma}^{-1} \sum_{i_2=1}^{n_1} \sum_{j_2=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_2}) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_2})^\top w_{i_2 j_2} \boldsymbol{\Sigma}^{-1} \\
& \times \frac{1}{n_2} \sum_{j_1=n_1+1}^{n} \varepsilon_{j_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_1}) \\
& + \boldsymbol{\gamma}_{1n}^\top \sum_{i=1}^{n_1} \sum_{j_2=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_2})^\top w_{i j_2} \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j_1=n_1+1}^{n} \varepsilon_{j_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_1}) \\
& + \frac{1}{n_1} \sum_{i_1=1}^{n_1} \varepsilon_{i_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_1})^\top \boldsymbol{\Sigma}^{-1} \sum_{i_2=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_2}) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{i_2 j} \delta_n \boldsymbol{\gamma}_{2n} \\
& + \boldsymbol{\gamma}_{1n}^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \boldsymbol{\gamma}_{2n} \\
& + \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \sum_{i=1}^{n_1} \sum_{j_2=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_2})^\top w_{i j_2} \boldsymbol{\Sigma}^{-1} \\
& \times \frac{1}{n_2} \sum_{j_1=n_1+1}^{n} \varepsilon_{j_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_1}) \\
& + \delta_n \frac{1}{n_1} \sum_{i_1=1}^{n_1} \varepsilon_{i_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_1})^\top \boldsymbol{\Sigma}^{-1} \sum_{i_2=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_2}) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{i_2 j} \boldsymbol{\Sigma}^{-1} \\
& \times E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
& + \delta_n \boldsymbol{\gamma}_{1n}^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
& + \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \boldsymbol{\gamma}_{2n} \\
& + \delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \boldsymbol{\Sigma}^{-1}
\end{aligned}
$$

$$\times E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}$$

$$=: \Gamma_{111} + \Gamma_{112} + \Gamma_{113} + \Gamma_{114} + \Gamma_{115} + \Gamma_{116} + \Gamma_{117} + \Gamma_{118} + \Gamma_{119}. \tag{S2.26}$$

Note that $\sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij}$ appears in every term in (S2.26). Hence, we analyse it for convenience in the future. Recall $\boldsymbol{\Sigma}_{cov}^* = E\left\{\dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_1\right) \dot{g}\left(\boldsymbol{\theta}^*, \boldsymbol{X}_2\right)^\top w_{12}\right\}$. Define

$$\boldsymbol{G}_n = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} - \boldsymbol{\Sigma}_{cov}^*. \tag{S2.27}$$

Then $\sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} = n_1 n_2 \boldsymbol{\Sigma}_{cov}^* + n_1 n_2 \boldsymbol{G}_n$. Under Condition 4, we have $\|\boldsymbol{\Sigma}_{cov}^*\| = O(1)$. Now we prove

$$\|\boldsymbol{G}_n\| = O_p\left(\max\left\{p/\sqrt{n_1}, p/\sqrt{n_2}\right\}\right) = o_p\left(1/\sqrt{p \log(\min\{n_1, n_2\})}\right).$$

Let

$$G_{kl}(\boldsymbol{X}_i, \boldsymbol{X}_j) := \dot{g}_k\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}_l(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{ij} - E\left\{\dot{g}_k\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) \dot{g}_l\left(\boldsymbol{\theta}_0, \boldsymbol{X}_2\right) w_{12}\right\}.$$

Then for any $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\sqrt{\frac{n_1 n_2}{(n_1+n_2)p^2}} \|\boldsymbol{G}_n\| > \epsilon\right) \\
\leq &\mathbb{P}\left(\frac{n_1 n_2}{(n_1+n_2)p^2} \sum_{k=1}^{p} \sum_{l=1}^{p} \left\{\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} G_{kl}(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\}^2 > \epsilon^2\right) \\
\leq &\frac{n_1 n_2}{(n_1+n_2)p^2} \sum_{k=1}^{p} \sum_{l=1}^{p} E\left[\left\{\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} G_{kl}(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\}^2\right] / \epsilon^2
\end{aligned}$$

$$\leq \frac{n_1 n_2}{(n_1+n_2)p^2} \sum_{k=1}^{p} \sum_{l=1}^{p} \frac{n_1+n_2-1}{n_1 n_2} E\left\{F^4(\boldsymbol{X})\right\}/\epsilon^2 = O(1), \qquad \text{(S2.28)}$$

where the last line is derived from the following illustration. By using Cauchy-Schwarz inequality, we have

$$\begin{aligned}
& E\left[\left\{\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} G_{kl}(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\}^2\right] \\
=& \frac{1}{n_1^2 n_2^2}\sum_{i_1=1}^{n_1}\sum_{i_2=1, i_2\neq i_1}^{n_1}\sum_{j=n_1+1}^{n} E\left\{G_{kl}(\boldsymbol{X}_{i_1}, \boldsymbol{X}_j) G_{kl}(\boldsymbol{X}_{i_2}, \boldsymbol{X}_j)\right\} \\
& + \frac{1}{n_1^2 n_2^2}\sum_{i=1}^{n_1}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1, j_2\neq j_1}^{n} E\left\{G_{kl}(\boldsymbol{X}_i, \boldsymbol{X}_{j_1}) G_{kl}(\boldsymbol{X}_i, \boldsymbol{X}_{j_2})\right\} \\
& + \frac{1}{n_1^2 n_2^2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} E\left\{G_{kl}^2(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\} \\
\leq& \frac{n_1-1}{n_1^2 n_2^2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} E\left\{G_{kl}^2(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\} + \frac{n_2-1}{n_1^2 n_2^2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} E\left\{G_{kl}^2(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\} \\
& + \frac{1}{n_1^2 n_2^2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} E\left\{G_{kl}^2(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\} \\
=& \frac{n_1+n_2-1}{n_1^2 n_2^2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} E\left\{G_{kl}^2(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\} \leq \frac{n_1+n_2-1}{n_1 n_2} E\left\{F^4(\boldsymbol{X})\right\},
\end{aligned}$$

where the last inequality holds because

$$\begin{aligned}
& E\left\{G_{kl}^2(\boldsymbol{X}_i, \boldsymbol{X}_j)\right\} \\
=& E\left(\left[\dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\, \dot{g}_l(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{ij} - E\left\{\dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\, \dot{g}_l(\boldsymbol{\theta}_0, \boldsymbol{X}_2)\, w_{12}\right\}\right]^2\right) \\
\leq& E\left[\left\{\dot{g}_k(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\, \dot{g}_l(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{ij}\right\}^2\right] \leq E\left\{F^4(\boldsymbol{X})\right\}.
\end{aligned}$$

Due to $p^3 \log n_1/n_1 \to 0$ and $p^3 \log n_2/n_2 \to 0$, from (S2.28), we obtain that

$$\|\boldsymbol{G}_n\| = O_p\left(\max\left\{p/\sqrt{n_1}, p/\sqrt{n_2}\right\}\right) = o_p\left(1/\sqrt{p\log(\min\{n_1,n_2\})}\right). \tag{S2.29}$$

Now we discuss the terms Γ_{11i}, $i = 1, 2, 3, \cdots, 9$ in (S2.26), respectively.

First, we proceed to prove

$$\Gamma_{111} = \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + o_p\left(\sqrt{n_1 n_2}\right). \tag{S2.30}$$

Decompose Γ_{111} as

$$\begin{aligned}
\Gamma_{111} =& \frac{1}{n_1} \sum_{i_1=1}^{n_1} \varepsilon_{i_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_1})^\top \boldsymbol{\Sigma}^{-1} \sum_{i_2=1}^{n_1} \sum_{j_2=n_1+1}^{n} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_2}\right) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_2})^\top w_{i_2 j_2} \boldsymbol{\Sigma}^{-1} \times \\
& \frac{1}{n_2} \sum_{j_1=n_1+1}^{n} \varepsilon_{j_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{j_1}) \\
=& \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
& + \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{G}_n \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
=:& \ \Gamma'_{111} + \Gamma''_{111}.
\end{aligned}$$

Next, we try to prove $\Gamma''_{111} = o_p\left(\sqrt{n_1 n_2}\right)$. Decompose Γ''_{111} as two terms. Let $\mathcal{D} := \{\|\sqrt{p\log(\min\{n_1,n_2\})}\boldsymbol{G}_n\| \le M_1\}$ and M_1 is a positive constant here. For any event $\mathcal{C}$, we let $\mathrm{I}_{\mathcal{C}}$ denote its indicator function and $\mathcal{C}^c$ its complement. Then $\Gamma''_{111} = \Gamma''_{111}\mathrm{I}_{\mathcal{D}} + \Gamma''_{111}\mathrm{I}_{\mathcal{D}^c}$. Recall

$\|\boldsymbol{G}_n\| = o_p\left(1/\sqrt{p\log(\min\{n_1,n_2\})}\right)$. Then $\mathbb{P}(\mathcal{D}) \to 1$, as $n_1, n_2 \to \infty$. It follows that ${\Gamma''}_{111}\mathrm{I}_{\mathcal{D}^c} = o_p(1)$. Therefore, it suffices to prove that ${\Gamma''}_{111}\mathrm{I}_{\mathcal{D}} = o_p(\sqrt{n_1 n_2})$, or equivalently, that $({\Gamma''}_{111})^2\mathrm{I}_{\mathcal{D}} = o_p(n_1 n_2)$. For ${\Gamma''}^2_{111}\mathrm{I}_{\mathcal{D}}$, we have

$$
\begin{aligned}
&E\left({\Gamma''}^2_{111}\mathrm{I}_{\mathcal{D}}\right)\\
=&E\left(\sum_{i=1}^{n_1}\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{G}_n\mathrm{I}_{\mathcal{D}}\boldsymbol{\Sigma}^{-1}\left\{\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\right\}\boldsymbol{\Sigma}^{-1}\right.\\
&\left.\quad\times\boldsymbol{G}_n\mathrm{I}_{\mathcal{D}}\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right)\\
\le&\left\|\boldsymbol{\Sigma}^{-1}\right\|^4 E\left(\left\|\sum_{i=1}^{n_1}\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\|^2\|\boldsymbol{G}_n\mathrm{I}_{\mathcal{D}}\|^2\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\right\|\right).
\end{aligned}
$$

From (S2.27), we know $\|\boldsymbol{G}_n\mathrm{I}_{\mathcal{D}}\|^2 \le M_1^2/\{p\log(\min\{n_1,n_2\})\}$. Notice that $\mathcal{N}_1$ and $\mathcal{N}_2$ are independent. Thus, we have

$$
\begin{aligned}
E\left({\Gamma''}^2_{111}\mathrm{I}_{\mathcal{D}}\right) \le&\left\|\boldsymbol{\Sigma}^{-1}\right\|^4 E\left\|\sum_{i=1}^{n_1}\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\|^2 E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\right\|\\
&\times M_1^2/\{p\log(\min\{n_1,n_2\})\}. \qquad \text{(S2.31)}
\end{aligned}
$$

Recall the definition of $\boldsymbol{\Sigma}_\varepsilon = E\left\{\varepsilon_1^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\right\}$ in Condition 4. Then $\left\|\boldsymbol{\Sigma}^{-1}\right\| = O(1)$ and $\|\boldsymbol{\Sigma}_\varepsilon\| = O(1)$. Note that

$$
\begin{aligned}
&E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\right\|^2\\
\le&E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top - n_2\boldsymbol{\Sigma}_\varepsilon\right\|^2 + n_2^2\|\boldsymbol{\Sigma}_\varepsilon\|^2
\end{aligned}
$$

$$\leq E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top-n_2\boldsymbol{\Sigma}_\varepsilon\right\|^2+O\left(n_2^2\right)=O\left(n_2^2\right), \quad \text{(S2.32)}$$

where the last equation is because if $p^2/n_2 \to 0$. Under Condition 8, we have

$$\begin{aligned}
&E\left\|\frac{1}{n_2}\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top-\boldsymbol{\Sigma}_\varepsilon\right\|^2\\
=&\sum_{k=1}^{p}\sum_{l=1}^{p}E\left[\frac{1}{n_2}\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}_k(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}_l(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-E\left\{\varepsilon_j^2\dot{g}_k(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}_l(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}\right]^2\\
=&\sum_{k=1}^{p}\sum_{l=1}^{p}\frac{1}{n_2}\operatorname{Var}\left\{\varepsilon_j^2\dot{g}_k(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}_l(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}\\
\leq&\sum_{k=1}^{p}\sum_{l=1}^{p}\frac{1}{n_2}E\left\{\varepsilon_j^4\dot{g}_k^2(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}_l^2(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}\\
=&O\left(p^2/n_2\right)=o_p\left(1\right),
\end{aligned}$$

then it follows that

$$E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top-n_2\boldsymbol{\Sigma}_\varepsilon\right\|^2=O\left(p^2n_2\right)=o\left(n_2^2\right). \quad \text{(S2.33)}$$

Thus, from (S2.32) and Lyapunov inequality, we gain

$$E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\right\|=O\left(n_2\right). \quad \text{(S2.34)}$$

Besides, from (S2.33), we know

$$E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top-n_2\boldsymbol{\Sigma}_\varepsilon\right\|=O\left(\sqrt{n_2}p\right), \quad \text{(S2.35)}$$

which will be useful when we discuss Γ_{112}. For $E\left\|\sum_{i=1}^{n_1}\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\|^2$ in

(S2.31), under Condition 2, using Cauchy-Schwarz inequality,

$$E\left\|\sum_{i=1}^{n_1}\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\|^2 = n_1\sum_{k=1}^{p}E\left\{\varepsilon_1^2\dot{g}_k^2(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\} \tag{S2.36}$$

$$\leq n_1 p E^{1/2}\left(\varepsilon_1^4\right)E^{1/2}\left\{F^4(\boldsymbol{X}_1)\right\} = O\left(n_1 p\right). \tag{S2.37}$$

Thus, from (S2.31) we know

$$E\left(\Gamma''^2_{111}\mathrm{I}_{\mathcal{D}}\right) = O\left(1\right)O\left(n_1 p\right)O\left(n_2\right)M_1^2/\{p\log(\min\{n_1,n_2\})\} = o\left(n_1 n_2\right).$$

Then $\Gamma''_{111} = o_p\left(\sqrt{n_1 n_2}\right)$ and so is $\Gamma_{111} = o_p\left(\sqrt{n_1 n_2}\right)$ in (S2.30).

Second, we prove $\Gamma_{11i} = o_p\left(\sqrt{n_1 n_2}\right)$ for $i = 2, 3, 4$. Decompose Γ_{112} as

$$\begin{aligned}\Gamma_{112} =& \boldsymbol{\gamma}_{1n}^\top\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_2})^\top w_{ij_2}\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j_1=n_1+1}^{n}\varepsilon_{j_1}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_1})\\ =& n_1\boldsymbol{\gamma}_{1n}^\top\boldsymbol{\Sigma}_{cov}^*\boldsymbol{\Sigma}^{-1}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j) + n_1\boldsymbol{\gamma}_{1n}^\top\boldsymbol{G}_n\boldsymbol{\Sigma}^{-1}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\\ =:& \Gamma'_{112}+\Gamma''_{112}.\end{aligned}$$

We now prove $E_{\mathcal{N}_1}\left(\Gamma'^2_{112}\right) = o_p\left(n_1 n_2\right)$, such that $\Gamma'_{112} = o_p\left(\sqrt{n_1 n_2}\right)$ by Markov inequality. Under Condition 4, $\|\boldsymbol{\Sigma}_{cov}^*\| = O_p\left(1\right)$ and $\left\|\boldsymbol{\Sigma}^{-1}\right\| = O_p\left(1\right)$, then we have

$$\begin{aligned}&E_{\mathcal{N}_1}\left(\Gamma'^2_{112}\right)\\ =& n_1^2\boldsymbol{\gamma}_{1n}^\top E_{\mathcal{N}_1}\left\{\boldsymbol{\Sigma}_{cov}^*\boldsymbol{\Sigma}^{-1}\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{cov}^*\right\}\boldsymbol{\gamma}_{1n}\\ \leq& n_1^2\|\boldsymbol{\gamma}_{1n}\|^2\left\|\boldsymbol{\Sigma}_{cov}^*\right\|^2\left\|\boldsymbol{\Sigma}^{-1}\right\|^2 E\left\|\sum_{j=n_1+1}^{n}\varepsilon_j^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\right\|\end{aligned}$$

$$
\begin{aligned}
\leq & n_1^2 \|\boldsymbol{\gamma}_{1n}\|^2 \left\|\boldsymbol{\Sigma}_{cov}^*\right\|^2 \left\|\boldsymbol{\Sigma}^{-1}\right\|^2 E \left\| \sum_{j=n_1+1}^{n} \varepsilon_j^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top - n_2 \boldsymbol{\Sigma}_\varepsilon \right\| \\
& + n_1^2 \|\boldsymbol{\gamma}_{1n}\|^2 \left\|\boldsymbol{\Sigma}_{cov}^*\right\|^2 \left\|\boldsymbol{\Sigma}^{-1}\right\|^2 n_2 \left\|\boldsymbol{\Sigma}_\varepsilon\right\| \\
= & n_1^2 o_p\left(1/n_1\right) E \left\| \sum_{j=n_1+1}^{n} \varepsilon_j^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top - n_2 \boldsymbol{\Sigma}_\varepsilon \right\| + n_1^2 n_2 o_p\left(1/n_1\right).
\end{aligned}
$$

Using equation (S2.35), we conclude that

$$
E_{\mathcal{N}_1}\left(\Gamma'^2_{112}\right) = n_1^2 o_p\left(1/n_1\right) O\left(\sqrt{n_2} p\right) + n_1^2 n_2 o_p\left(1/n_1\right) = o_p\left(n_1 n_2\right),
$$

so $\Gamma'_{112} = o_p\left(\sqrt{n_1 n_2}\right)$. Recall $\|\boldsymbol{\gamma}_{1n}\| = o_p\left(1/\sqrt{n_1}\right)$. Note that $\Gamma''_{112} = \Gamma''_{112}\mathrm{I}_{\mathcal{D}} + \Gamma''_{112}\mathrm{I}_{\mathcal{D}^c}$, and

$$
\begin{aligned}
& E_{\mathcal{N}_1}\left(\Gamma''^2_{112}\mathrm{I}_{\mathcal{D}}\right) \\
= & n_1^2 \boldsymbol{\gamma}_{1n}^\top E_{\mathcal{N}_1} \left\{ \boldsymbol{G}_n \mathrm{I}_{\mathcal{D}} \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{G}_n \mathrm{I}_{\mathcal{D}} \right\} \boldsymbol{\gamma}_{1n} \\
\leq & n_1^2 \|\boldsymbol{\gamma}_{1n}\|^2 \|\boldsymbol{G}_n \mathrm{I}_{\mathcal{D}}\|^2 \left\|\boldsymbol{\Sigma}^{-1}\right\|^2 E_{\mathcal{N}_1} \left\{ \sum_{j=n_1+1}^{n} \varepsilon_j^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \right\} \\
= & n_1^2 \|\boldsymbol{\gamma}_{1n}\|^2 \|\boldsymbol{G}_n \mathrm{I}_{\mathcal{D}}\|^2 \left\|\boldsymbol{\Sigma}^{-1}\right\|^2 n_2 \left\|\boldsymbol{\Sigma}_\varepsilon\right\| \\
\leq & n_1^2 o_p\left(1/n_1\right) M_1^2 / \{p \log(\min\{n_1, n_2\})\} O\left(n_2\right) = o_p\left(n_1 n_2\right).
\end{aligned}
$$

Then we obtain $\Gamma''_{112} = o_p\left(\sqrt{n_1 n_2}\right)$. So $\Gamma_{112} = \Gamma'_{112} + \Gamma''_{112} = o_p\left(\sqrt{n_1 n_2}\right)$.

Since Γ_{113} has the same structure as Γ_{112}, we have $\Gamma_{113} = o_p\left(\sqrt{n_1 n_2}\right)$. For Γ_{114}, we have

$$
\Gamma_{114} = \boldsymbol{\gamma}_{1n}^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \boldsymbol{\gamma}_{2n}
$$

$$
\begin{aligned}
=&\boldsymbol{\gamma}_{1n}^{\top}\left\{\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}w_{ij}-n_1n_2\boldsymbol{\Sigma}_{\varepsilon}\right\}\boldsymbol{\gamma}_{2n}\\
&+\boldsymbol{\gamma}_{1n}^{\top}n_1n_2\boldsymbol{\Sigma}_{\varepsilon}\boldsymbol{\gamma}_{2n}\\
=&n_1n_2\boldsymbol{\gamma}_{1n}^{\top}\boldsymbol{G}_n\boldsymbol{\gamma}_{2n}+\boldsymbol{\gamma}_{1n}^{\top}n_1n_2\boldsymbol{\Sigma}_{\varepsilon}\boldsymbol{\gamma}_{2n}\\
\leq&n_1n_2\left\|\boldsymbol{\gamma}_{1n}\right\|\left\|\boldsymbol{G}_n\right\|\|\boldsymbol{\gamma}_{2n}\|+n_1n_2\|\boldsymbol{\gamma}_{1n}\|\|\boldsymbol{\Sigma}_{\varepsilon}\|\|\boldsymbol{\gamma}_{2n}\|=o_p\left(\sqrt{n_1n_2}\right).
\end{aligned}
$$

Third, we deal with Γ_{115} and Γ_{116}. Decompose Γ_{115} as

$$
\begin{aligned}
\Gamma_{115}=&\delta_nE\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}^{\top}\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n_1}\sum_{j_2=n_1+1}^{n}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_2})^{\top}w_{ij_2}\times\\
&\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j_1=n_1+1}^{n}\varepsilon_{j_1}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_1})\\
=&n_1n_2\delta_nE\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{cov}^{*}\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\\
&+n_1n_2\delta_nE\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{G}_n\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j).
\end{aligned}
\tag{S2.38}
$$

Analyse the order of $\|E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}\|$ first. Under Condition 10, $\left\|E\left\{l^2(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\right\}\right\|=O\left(1\right)$. Then we have that

$$
\begin{aligned}
&\|E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\}\|^2\\
=&\sup_{\|\boldsymbol{a}\|=1}\left[\boldsymbol{a}^{\top}E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\}\right]^2\\
=&\sup_{\|\boldsymbol{a}\|=1}E^2\left\{l(\boldsymbol{X}_1)\boldsymbol{a}^{\top}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\}\\
\leq&\sup_{\|\boldsymbol{a}\|=1}E\left\{l^2(\boldsymbol{X}_1)\boldsymbol{a}^{\top}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{a}\right\}
\end{aligned}
$$

$$= \left\| E\left\{ l^2(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \right\} \right\| = O(1). \tag{S2.39}$$

From (S2.36), we derive that $\frac{1}{n_2}\sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) = O_p\left(\sqrt{p/n_2}\right)$. Recall $\|\boldsymbol{G}_n\| = O_p\left(\max\left\{p/\sqrt{n_1}, p/\sqrt{n_2}\right\}\right)$. Thus,

$$\begin{aligned} & n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{G}_n \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\ =& n_1 n_2 \delta_n O(1) O_p\left(\max\left\{p/\sqrt{n_1}, p/\sqrt{n_2}\right\}\right) O_p\left(\sqrt{p/n_2}\right) = o_p\left(n_1\sqrt{n_2}\delta_n\right). \end{aligned}$$

From (S2.38), we know

$$\begin{aligned} \Gamma_{115} =& n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\ & + o_p\left(n_1\sqrt{n_2}\delta_n\right). \end{aligned}$$

Similarly, we have

$$\begin{aligned} \Gamma_{116} =& n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \\ & + o_p\left(\sqrt{n_1} n_2 \delta_n\right). \end{aligned}$$

Fourth, we try to prove $\Gamma_{11i} = o_p\left(\sqrt{n_1 n_2}\right)$ for $i = 7, 8$ and

$$\begin{aligned} \Gamma_{119} =& n_1 n_2 \delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\ & + o_p\left(n_1 n_2 \delta_n^2\right). \end{aligned} \tag{S2.40}$$

Similar to the arguments used for analyzing Γ_{111} to Γ_{116}, we have $\Gamma_{117} =$

$o_p\left(\sqrt{n_1}n_2\delta_n\right)$ and $\Gamma_{118} = o_p\left(n_1\sqrt{n_2}\delta_n\right)$. For Γ_{119}, we have

$$\begin{aligned}\Gamma_{119} =&\delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \boldsymbol{\Sigma}^{-1} \\ &\times E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\ =&n_1 n_2 \delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\ &+ o_p\left(n_1 n_2 \delta_n^2\right).\end{aligned}$$

Altogether, we can summarize the results to achieve the decomposition as,

$$\begin{aligned}\Gamma_{11} =&\Gamma_{111} + \Gamma_{115} + \Gamma_{116} + \Gamma_{119} \\ &+ o_p\left(\max\left\{\sqrt{n_1 n_2}, \sqrt{n_1} n_2 \delta_n, n_1\sqrt{n_2}\delta_n, n_1 n_2 \delta_n^2\right\}\right) \\ =&\sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + o_p\left(\sqrt{n_1 n_2}\right) \\ &+ n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\ &+ n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \\ &+ n_1 n_2 \delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\ &+ o_p\left(\max\left\{\sqrt{n_1 n_2}, \sqrt{n_1} n_2 \delta_n, n_1\sqrt{n_2}\delta_n, n_1 n_2 \delta_n^2\right\}\right).\end{aligned}$$

The proof of (S2.23) is finished. That is the end of *Step 1.*

Step 2. Below we prove $\Gamma_{12} = o_p\left(\max\left\{n_1\sqrt{n_2}\delta_n, \sqrt{n_1 n_2}\right\}\right)$. To take the advantage of the result $\left\|E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{1j}\right\}\right\| \leq \sqrt{b^* E_{\boldsymbol{X}_j}\left(w_{1j}^4\right)}$ in Remark 1, we decompose $\sum_{i=1}^{n_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top w_{ij}$ in Γ_{12} as the sum

of $n_2 \left\| E_{\boldsymbol{X}_j} \{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j}\} \right\|$ and $n_2 \boldsymbol{S}_{jn_1}$, where we define $\boldsymbol{S}_{jn_1} = \frac{1}{n_1} \sum_{i=1}^{n_1} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i) w_{ij} - E_{\boldsymbol{X}_j} \{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j}\}$. Then it follows that

$$\begin{aligned}
\Gamma_{12} =& \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top w_{ij} \\
& \times \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
=& \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top n_1 \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_j} \{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\} \\
& \times \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_i\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
& + \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top n_1 \sum_{j=n_1+1}^{n} \boldsymbol{S}_{jn_1} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
=:& \Gamma_{121} + \Gamma_{122}. \qquad \text{(S2.41)}
\end{aligned}$$

First, analyse Γ_{121}.

$$\begin{aligned}
\Gamma_{121} =& \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_j} \{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\} \\
& \times \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
& + n_1 \delta_n E \{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\}^\top \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_j} \{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\} \\
& \times \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
& + \boldsymbol{\gamma}_{1n}^\top \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_j} \{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\} \\
& \times \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
=:& \Gamma'_{121} + \Gamma''_{121} + \Gamma'''_{121}. \qquad \text{(S2.42)}
\end{aligned}$$

We now prove $E_{\mathcal{N}_2}\left(\Gamma'^2_{121}\right) = o_p\left(n_1 n_2\right)$ in the following, which leads to $\Gamma'_{121} = o_p\left(\sqrt{n_1 n_2}\right)$. When $i_1 \neq i_2$, $E\left\{\varepsilon_{i_1}\varepsilon_{i_2}\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_1})\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_{i_2})^\top\right\} = 0$. Thus,

$$
\begin{aligned}
&E_{\mathcal{N}_2}\left(\Gamma'^2_{121}\right) \\
=&\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{1j}\right\}^\top \boldsymbol{\Sigma}^{-1} \\
&\quad \times n_1 E\left\{\sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top\right\} \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} \\
&\quad \times \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
=&\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{1j}\right\}^\top \boldsymbol{\Sigma}^{-1} \\
&\quad \times n_1 \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} \\
&\quad \times \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
\leq& n_1 \left\|\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{1j}\right\}^\top\right\|^2 \\
&\quad \times \left\|\boldsymbol{\Sigma}^{-1}\right\|^2 \left\|\boldsymbol{\Sigma}_\varepsilon\right\|, \qquad \text{(S2.43)}
\end{aligned}
$$

where $\boldsymbol{\Sigma}_\varepsilon = E\left\{\varepsilon_1^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top\right\}$. Under Condition 6, since $\tilde{\boldsymbol{\theta}}_2^*$ lies between $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}_2$, we have

$$
\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\| \leq \sqrt{p}\left\|\tilde{\boldsymbol{\theta}}_2^* - \boldsymbol{\theta}_0\right\| L(\boldsymbol{X}_j) \leq \sqrt{p}\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\| L(\boldsymbol{X}_j) \qquad \text{(S2.44)}
$$

and $E\{L^4(\boldsymbol{X}_1)\} = O(1)$. Recall $\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\| = O_p\left(\max\left\{\sqrt{p/n_2}, \sqrt{p}\delta_n\right\}\right)$ in Lemma 2, $p^3 \log n_2/n_2 \to 0$. Thus,

$$\begin{aligned}
&\left\|\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}^\top\right\| \\
\leq & \sum_{j=n_1+1}^{n} \left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\| \left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\| \left\|E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}\right\| \\
\leq & \left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\|^2 \sum_{j=n_1+1}^{n} \sqrt{p} L(\boldsymbol{X}_j) \sqrt{b^* E_{\boldsymbol{X}_1}\left(w_{12}^2\right)} = o_p\left(\sqrt{n_2}\right). \qquad \text{(S2.45)}
\end{aligned}$$

Under Condition 4, $\left\|\boldsymbol{\Sigma}^{-1}\right\|^2 \|\boldsymbol{\Sigma}_\varepsilon\| = O(1)$. Thus, from (S2.43), we know $E_{\mathcal{N}_2}\left(\Gamma'^2_{121}\right) = o_p(n_1 n_2)$. The application of Markov inequality yields $\Gamma'_{121} = o_p\left(\sqrt{n_1 n_2}\right)$. Note that the only difference between Γ'_{121} and Γ''_{121} is that we replace $\sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)$ with $n_1 \delta_n E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\}$. Taking the advantage of $\|E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\}\| = O(1)$ in (S2.39), we can then derive $\Gamma''_{121} = o_p\left(\max\left\{\sqrt{n_1 n_2}, n_1\sqrt{n_2}\delta_n\right\}\right)$. Omit the details here. Similarly, we have $\Gamma'''_{121} = o_p\left(\sqrt{n_1 n_2}\right)$.

Hence, $\Gamma_{121} = o_p\left(\max\left\{\sqrt{n_1 n_2}, n_1\sqrt{n_2}\delta_n\right\}\right)$.

Second, analyse Γ_{122}. Recall

$$\Gamma_{122} = \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top n_1 \sum_{j=n_1+1}^{n} \boldsymbol{S}_{jn_1} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right), \qquad \text{(S2.46)}$$

where $\boldsymbol{S}_{jn_1} = \frac{1}{n_1}\sum_{i=1}^{n_1} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i) w_{ij} - E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{1j}\}$. According to

Cauchy-Schwarz inequality,

$$\begin{aligned}&\left\|\sum_{j=n_1+1}^{n}\boldsymbol{S}_{jn_1}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*,\boldsymbol{X}_j\right)-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}^\top\right\|\\ \leq&\left(\sum_{j=n_1+1}^{n}\|\boldsymbol{S}_{jn_1}\|^2\right)^{1/2}\left\{\sum_{j=n_1+1}^{n}\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*,\boldsymbol{X}_j\right)-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^2\right\}^{1/2}. \quad \text{(S2.47)}\end{aligned}$$

Similar to (S2.13), we have $E\left(\|\boldsymbol{S}_{jn_1}\|^4\right) \leq \frac{p^2}{n_2^2}E\left\{F^4(\boldsymbol{X}_1)\right\} = O\left(p^2/n_2^2\right)$. Thus, $\left(\sum_{j=n_1+1}^{n}\|\boldsymbol{S}_{jn_1}\|^2\right)^{1/2} = O_p\left(\sqrt{p}\right)$.

Recall $\left\|\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right\| = O_p\left(\max\left\{\sqrt{p/n_2},\sqrt{p}\delta_n\right\}\right)$. From (S2.44) we know $\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*,\boldsymbol{X}_j\right)-\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)\right\| \leq \sqrt{p}\left\|\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right\|L(\boldsymbol{X}_j)$. Then it follows that

$$\begin{aligned}\left\{\sum_{j=n_1+1}^{n}\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*,\boldsymbol{X}_j\right)-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^2\right\}^{1/2} \leq&\left\{\sum_{j=n_1+1}^{n}p\left\|\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right\|^2L^2(\boldsymbol{X}_j)\right\}^{1/2}\\ =&O_p\left(\max\left\{p,p\sqrt{n_2}\delta_n\right\}\right). \quad \text{(S2.48)}\end{aligned}$$

Combining (S2.46), (S2.47) and (S2.48), we obtain

$$\begin{aligned}|\Gamma_{122}| \leq&n_1\left\|\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right\|\left(\sum_{j=n_1+1}^{n}\|\boldsymbol{S}_{jn_1}\|^2\right)^{1/2}\\ &\times\left\{\sum_{j=n_1+1}^{n}\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_2^*,\boldsymbol{X}_j\right)-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\|^2\right\}^{1/2}\left\|\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right\|\\ =&n_1O_p\left(\max\left\{\sqrt{p/n_1},\sqrt{p}\delta_n\right\}\right)O_p\left(\sqrt{p}\right)O_p\left(\max\left\{p,p\sqrt{n_2}\delta_n\right\}\right)\\ &\times O_p\left(\max\left\{\sqrt{p/n_2},\sqrt{p}\delta_n\right\}\right)\\ =&O_p\left(\max\left\{p^{5/2}\sqrt{n_1/n_2},p^{5/2}\sqrt{n_1n_2}\delta_n^2,p^{5/2}\delta_nn_1/\sqrt{n_2},p^{5/2}n_1\sqrt{n_2}\delta_n^3\right\}\right).\end{aligned}$$

Since $p^3\log n_1/n_1 \to 0$ and $p^3\log n_2/n_2 \to 0$, $\Gamma_{122} = o_p\left(\max\left\{n_1\sqrt{n_2}\delta_n,\sqrt{n_1n_2}\right\}\right)$.

Then we derive $\Gamma_{12} = o_p\left(\max\left\{n_1\sqrt{n_2}\delta_n, \sqrt{n_1 n_2}\right\}\right)$ from (S2.41).

That is the end of *Step 2.*

Step 3. We try to prove $\Gamma_2 = o_p\left(\max\left\{\sqrt{n_1 n_2}, \sqrt{n_1} n_2 \delta_n\right\}\right)$. Γ_2 can be rewritten as

$$\begin{aligned}
\Gamma_2 &= \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_1^*, \boldsymbol{X}_i\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\right\} \\
&\qquad \times \left\{g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) - g(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\} w_{ij} \\
&= \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_1^*, \boldsymbol{X}_i\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\right\} \\
&\qquad \times \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top w_{ij} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
&\quad + \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_1^*, \boldsymbol{X}_i\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\right\} \\
&\qquad \times \left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_2'^*, \boldsymbol{X}_j\right) - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^\top w_{ij} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\
&=: \Gamma_{21} + \Gamma_{22},
\end{aligned}$$

where $\tilde{\boldsymbol{\theta}}_1^*$ lies between $\hat{\boldsymbol{\theta}}_1$ and $\boldsymbol{\theta}_0$, and $\tilde{\boldsymbol{\theta}}_2'^*$ lies between $\hat{\boldsymbol{\theta}}_2$ and $\boldsymbol{\theta}_0$.

Similar to Γ_{12}, we have $\Gamma_{21} = o_p\left(\max\left\{\sqrt{n_1} n_2 \delta_n, \sqrt{n_1 n_2}\right\}\right)$. We prove $\Gamma_{22} = o_p\left(\sqrt{n_1 n_2}\right)$ in the following. According to (S2.44), $\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_1^*, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)\right\| \leq \sqrt{p}\left\|\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right\| L(\boldsymbol{X}_i)$ and

$$\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_2'^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\| \leq \sqrt{p}\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\| L(\boldsymbol{X}_j).$$

Thus,

$$
\begin{aligned}
|\Gamma_{22}| \leq & \left\|\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right\| \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n}\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_1^*, \boldsymbol{X}_i\right)-\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\right\| \\
& \times\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}'^*_2, \boldsymbol{X}_j\right)-\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\| w_{ij}\left\|\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right\|^2 \\
\leq & p\left\|\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right\|^2\left\|\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right\|^2 \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} L(\boldsymbol{X}_i) L(\boldsymbol{X}_j) w_{ij} \\
= & O_p\left(\max\left\{p/n_1, p\delta_n^2\right\}\right) O_p\left(\max\left\{p/n_2, p\delta_n^2\right\}\right) O\left(n_1 n_2\right) \\
= & O_p\left(\max\left\{p^3, p^3 n_1\delta_n^2, p^3 n_2\delta_n^2, p^3 n_1 n_2\delta_n^4\right\}\right)=o_p\left(\max\{\sqrt{n_1 n_2}, n_1 n_2\delta_n^2\}\right),
\end{aligned}
$$

where the last equation is derived from $p^3\delta_n^2 \to 0$, $p^3 \log n_1/n_1 \to 0$ and $p^3 \log n_2/n_2 \to 0$.

So $\Gamma_2 = o_p\left(\max\left\{\sqrt{n_1 n_2}, \sqrt{n_1} n_2\delta_n\right\}\right)$.

Now we finish *Step 1-3* to achieve

$$
\begin{aligned}
& \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
= & \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
& + n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
& + n_1 n_2 \delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \\
& + n_1 n_2 \delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
& + o_p\left(\max\left\{\sqrt{n_1 n_2}, \sqrt{n_1} n_2\delta_n, n_1\sqrt{n_2}\delta_n, n_1 n_2\delta_n^2\right\}\right).
\end{aligned}
$$

The conclusion of case (a) holds with $\delta_n = 0$, and thus the proof is finished.

□

The statistic before the studentization has the form of

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}.$$

We concern if we replace $\hat{e}_i$, $\hat{e}_j$ with ε_i, ε_j, how close the two terms will be. Recall that

$$\begin{aligned} w'_{ij} =& w_{ij} - \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_i} \{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) w_{ij}\} \\ & - \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i) w_{ij}\} \\ & + \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j). \end{aligned}$$

Lemma 5 shows the relationship between $\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}$ and $\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij}$ under the null hypothesis. For local alternative hypothesis, see Lemma 12.

Lemma 5. *Under the null hypothesis in (S2.6), when Conditions 1-8 hold, we have*

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} = \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + o_p(1). \quad \text{(S2.49)}$$

Proof of Lemma 5. Using case (a) in Lemma 3 and case (a) in Lemma 4, we conclude that

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_2+1}^{n} \hat{e}_i \hat{e}_j w_{ij}$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w_{ij}\\
&+\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\\
&+\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\varepsilon_j w_{ij}\\
&+\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\\
&\quad\times\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\\
=&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w_{ij}\\
&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{i2}\right\}\\
&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_i\right)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)w_{1j}\right\}\\
&+\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{*}_{cov}\boldsymbol{\Sigma}^{-1}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)+o_p\left(1\right)\\
=&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1=1}^{n}\varepsilon_i\varepsilon_j w'_{ij}+o_p\left(1\right).
\end{aligned}
\tag{S2.50}
$$

Here, in $E_{\boldsymbol{X}_i}\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{i2}\}$, we define "$\boldsymbol{X}_2$" as a copy of $\boldsymbol{X}$ which is independent of $\boldsymbol{X}_i$. The proof is finished. □

Recall that

$$
\tilde{w}_i(\mathcal{N}_2)=\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j\left[w_{ij}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}\right].
$$

The following lemma and Lemma 9 play important roles in explaining that

$$V_n = \frac{\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2)}{\sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \varepsilon_i \tilde{w}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2) \right\}^2}}$$

in our paper and

$$T_n := \frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij}}{E_{\mathcal{N}_2} \left\{ \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right)^2 \right\}}$$

have the same limiting null distribution, and we will show that T_n converges weakly to the standard normal distribution in the proof of **Theorem 1**.

Lemma 6. *If Conditions 1-8 hold, under the null hypothesis in (S2.6), we have*

$$\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + o_p(1). \tag{S2.51}$$

Proof of Lemma 6. To present the proof clearly, before our discussion about $\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2)$, we make some analysis on $\tilde{w}_1(\mathcal{N}_2)$. Noting that $\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j w_{1j}$ plays an important role in $\tilde{w}_1(\mathcal{N}_2)$, we make a discussion about $\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j w_{1j}$. Under the null hypothesis, from Lemma 3, we have

$$\begin{aligned} &\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j w_{1j} \\ =&\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ \varepsilon_j + g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j \right) \right\} w_{1j} \end{aligned} \tag{S2.52}$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{1j}\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}-R_1\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\right]-R_1, \quad \text{(S2.53)}
\end{aligned}
$$

where

$$
E_{\mathcal{N}_2}(R_1^4)=o_p\left(1\right). \quad \text{(S2.54)}
$$

Next, we analyse $\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\}$ in $\tilde{w}_1(\mathcal{N}_2)$. For this term, using equation (S2.52),

$$
\begin{aligned}
&\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\}\\
=&\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j w_{1j}\right\}\\
=&\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\mathcal{N}_2}\Bigg\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}\right.\\
&\left.\times E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\right]\Bigg\}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)R_1\right\}.
\end{aligned}
\quad \text{(S2.55)}
$$

Now we analyse the second term of equation (S2.55). Firstly, we illustrate that

$$
\left\|E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)R_1\right\}\right\|=o_p\left(1\right). \quad \text{(S2.56)}
$$

Under Lemma 3, $E_{\mathcal{N}_2}\left(R_1^4\right)=o_p\left(1\right)$. Thus, using Cauchy-Schwarz inequal-

ity, we can derive that under Condition 4,

$$
\begin{aligned}
&\left\| E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}\right\|^2 \\
=& \sup_{\|\boldsymbol{a}\|=1}\left[\boldsymbol{a}^{\top} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}\right]^2 \\
=& \sup_{\|\boldsymbol{a}\|=1} E_{\mathcal{N}_2}^2\left\{\boldsymbol{a}^{\top} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\} \\
\leq& \sup_{\|\boldsymbol{a}\|=1} E_{\mathcal{N}_2}\left\{\left|\boldsymbol{a}^{\top} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\right|^2\right\} E_{\mathcal{N}_2}\left(R_1^2\right) \\
=& \sup_{\|\boldsymbol{a}\|=1} E_{\mathcal{N}_2}\left\{\boldsymbol{a}^{\top} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^{\top} \boldsymbol{a}\right\} E_{\mathcal{N}_2}\left(R_1^2\right) \\
=& \left\| E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^{\top}\right\}\right\| E_{\mathcal{N}_2}\left(R_1^2\right) \\
=& \|\boldsymbol{\Sigma}\| E_{\mathcal{N}_2}\left(R_1^2\right) \leq b E_{\mathcal{N}_2}\left(R_1^2\right)=o_p(1) .
\end{aligned} \tag{S2.57}
$$

This implies that

$$
\begin{aligned}
&E_{\mathcal{N}_2}\left[\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^{\top} \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}\right]^2 \\
=& E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}^{\top} \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^{\top}\right\} \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\} \\
=& E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}^{\top} \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}=o_p(1) .
\end{aligned} \tag{S2.58}
$$

Then Markov inequality yields that

$$
\tilde{R}_1 := \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right)^{\top} \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}=o_p(1) . \tag{S2.59}
$$

According to (S2.52), (S2.55) and (S2.59), $\tilde{w}_1(\mathcal{N}_2)$ can be written as follows:

$$
\tilde{w}_1(\mathcal{N}_2) \tag{S2.60}
$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\}\right]\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\}\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\right]\\
&-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\\
&\quad\times E_{\boldsymbol{X}_j}\left[\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\left\{w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\right]-R_1+\tilde{R}_1\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\right\}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)^\top w_{12}\right\}\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\\
&-R_1+\tilde{R}_1\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}-R_1+\tilde{R}_1. \qquad \text{(S2.61)}
\end{aligned}
$$

The last equation follows from the definition of w'_{1j} in equation (S1.1) before Condition 9.

Now we start to study $\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)$. Let $R_1^*:=-R_1+\tilde{R}_1$. Similarly, we can define R_i^* for $i\neq 1$. Then $\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)$ can be rewritten as

$$
\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)=\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}+\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\varepsilon_i R_i^*
$$

$$=: \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + R_{(1)}.$$

Now we prove that $R_{(1)} = o_p(1)$. Notice that

$$\begin{aligned} E_{\mathcal{N}_2}\left(R_{(1)}^2\right) = \frac{1}{n_1} \sum_{i=1}^{n_1} E_{\mathcal{N}_2}\left(\varepsilon_i^2 R_i^{*2}\right) =& E_{\mathcal{N}_2}\left(\varepsilon_1^2 R_1^{*2}\right) \\ \leq& 2E_{\mathcal{N}_2}\left(\varepsilon_1^2 R_1^2\right) + 2E_{\mathcal{N}_2}\left(\varepsilon_1^2 \tilde{R}_1^2\right). \end{aligned}$$

We discuss $E_{\mathcal{N}_2}\left(\varepsilon_1^2 R_1^2\right)$ and $E_{\mathcal{N}_2}\left(\varepsilon_1^2 \tilde{R}_1^2\right)$ in the following. From equation (S2.54), we know $E_{\mathcal{N}_2}\left(R_1^4\right) = o_p(1)$. Then using Cauchy-Schwarz inequality, we conclude that

$$E_{\mathcal{N}_2}\left(\varepsilon_1^2 R_1^2\right) \leq E_{\mathcal{N}_2}^{1/2}\left(\varepsilon_1^4\right) E_{\mathcal{N}_2}^{1/2}\left(R_1^4\right) = o_p(1).$$

And for $E_{\mathcal{N}_2}\left(\varepsilon_1^2 \tilde{R}_1^2\right)$, using (S2.56), we have

$$\begin{aligned} & E_{\mathcal{N}_2}\left(\varepsilon_1^2 \tilde{R}_1^2\right) \\ =& E_{\mathcal{N}_2}\left(\varepsilon_1^2 E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}^\top \boldsymbol{\Sigma}^{-1} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right)^\top \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}\right) \\ =& E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\varepsilon_1^2 \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right)^\top\right\} \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\} \\ =& E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\} = o_p(1), \end{aligned}$$

where $\boldsymbol{\Sigma}_\varepsilon = E\left\{\varepsilon_1^2 \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right)^\top\right\}$. Hence,

$$E_{\mathcal{N}_2}\left(R_{(1)}^2\right) = E_{\mathcal{N}_2}\left(\varepsilon_1^2 R_1^{*2}\right) = o_p(1). \tag{S2.62}$$

Using Markov inequality, we know $R_{(1)} = o_p(1)$. Then it follows that

$$\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + o_p(1).$$

The proof of (S2.51) is finished. □

Then we have the following corollary, equation (4) in the main context.

Corollary 1. *Suppose Conditions 1-8 hold. Under the null hypothesis in (S2.6),*

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} = \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} e_i \tilde{w}_i(\mathcal{N}_2) + o_p(1).$$

Proof of Corollary 1. Under the null hypothesis in (S2.6), recall $\varepsilon_i = e_i$. Then Corollary 1 is a direct result of Lemma 5 and 6. □

Now we use the martingale central limit theorem to prove two lemmas, which will be useful in the future. The following lemma is useful for Lemma 13. In this lemma, we denote $E_{\boldsymbol{X}_1=\boldsymbol{x}}(\cdot) = E(\cdot|\boldsymbol{X}_1 = \boldsymbol{x})$ as the conditional expectation over other random variable(s) given $\boldsymbol{X}_1 = \boldsymbol{x}$.

Lemma 7. *Under Condition 9, given $\boldsymbol{X}_1 = \boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^q$, we have*

$$\left. \frac{\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j w'_{1j}}{E^{1/2}_{\boldsymbol{X}_1=\boldsymbol{x}} \left(\varepsilon_2^2 w'^2_{12}\right)} \right| \{\boldsymbol{X}_1 = \boldsymbol{x}\} \xrightarrow{d} \mathbf{N}(0,1).$$

Proof of Lemma 7. We use the martingale central limit theorem to prove it. Define $\eta_{nj} = \frac{\frac{1}{\sqrt{n_2}} \varepsilon_j w'_{1j}}{E^{1/2}_{\boldsymbol{X}_1=\boldsymbol{x}} \left(\varepsilon_2^2 w'^2_{12}\right)}$ and

$$\mathcal{F}_{n,j} = \sigma\left\{\left(\boldsymbol{X}^\top_{n_1+1}, \varepsilon_{n_1+1}\right), \left(\boldsymbol{X}^\top_{n_1+2}, \varepsilon_{n_1+2}\right), \cdots, \left(\boldsymbol{X}^\top_{j}, \varepsilon_{j}\right)\right\}, n_1 + 1 \le j \le n,$$

which is the σ algebra generated by

$$\left\{\left(\boldsymbol{X}^\top_{n_1+1}, \varepsilon_{n_1+1}\right), \left(\boldsymbol{X}^\top_{n_1+2}, \varepsilon_{n_1+2}\right), \cdots, \left(\boldsymbol{X}^\top_{j}, \varepsilon_{j}\right)\right\}.$$

Besides, let $\mathcal{F}_{n,n_1} = \emptyset$. It is easy to check $E_{\boldsymbol{X}_1=\boldsymbol{x}}(\eta_{nj}|\mathcal{F}_{n,j-1}) = 0$ and $\left\{\sum_{j=n_1+1}^{m}\eta_{nj}, \mathcal{F}_{n,m}\right\}_{m=n_1+1}^{n}$ is a zero mean martingle. Because

$$\begin{aligned}
E_{\boldsymbol{X}_1=\boldsymbol{x}}\left\{\left(\sum_{j=n_1+1}^{m}\eta_{nj}\right)^2\right\} =& E_{\boldsymbol{X}_1=\boldsymbol{x}}\left\{\left(\frac{\sum_{j=n_1+1}^{m}\frac{1}{\sqrt{n_2}}\varepsilon_j w'_{1j}}{E^{1/2}_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)}\right)^2\right\} \\
=& \frac{E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\sum_{j=n_1+1}^{m}\frac{1}{\sqrt{n_2}}\varepsilon_j w'_{1j}\right)^2}{E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)} \\
=& \frac{1}{n_2}\sum_{j=n_1+1}^{m}\frac{E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)}{E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)} = (m-n_1)/n_2 \leq 1,
\end{aligned}$$

this martingale is square integrable. The martingale central limit theorem in Hall and Heyde (1980) holds if we can show that

$$\sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\eta_{nj}^2|\mathcal{F}_{n,j-1}\right) \xrightarrow{p} 1 \tag{S2.63}$$

and for any $\epsilon > 0$,

$$V_{nn_2} = \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\eta_{nj}^2 \mathrm{I}_{\{\eta_{nj}>\epsilon\}}|\mathcal{F}_{n,j-1}\right) \xrightarrow{p} 0. \tag{S2.64}$$

Clearly,

$$\sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\eta_{nj}^2|\mathcal{F}_{n,j-1}\right) = \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\frac{\frac{1}{n_2}\varepsilon_j^2 w'^2_{1j}}{E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)}\middle|\mathcal{F}_{n,j-1}\right) = 1.$$

So $(S2.63)$ holds. We prove (S2.64) in the following. For V_{nn_2}, noting that η_{nj}^2 is independent of $\mathcal{F}_{n,j-1}$, we infer that

$$\begin{aligned}
V_{nn_2} =& \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\eta_{nj}^2 \mathrm{I}_{\{\eta_{nj}>\epsilon\}}|\mathcal{F}_{n,j-1}\right) \\
=& \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\eta_{nj}^2 \mathrm{I}_{\{\eta_{nj}>\epsilon\}}\right)
\end{aligned}$$

$$=E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}}\right). \tag{S2.65}$$

Here we have $\left|n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}}\right| \leq n_2\eta_{n2}^2 = \frac{\varepsilon_2^2 w'^2_{12}}{E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)}$ and $E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\frac{\varepsilon_2^2 w'^2_{12}}{E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)}\right) = 1$. Thus, $n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}}$ is uniformly integrable. If we can derive that $n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}} = o_p(1)$, then $E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}}\right) = o(1)$. For any $\tau > 0$, noting that $n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}} = 0$ if $\eta_{n2} \leq \epsilon$, we have

$$\begin{aligned}
&\mathbb{P}\left(n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}} \geq \tau \middle| \{\boldsymbol{X}_1 = \boldsymbol{x}\}\right) \leq \mathbb{P}\left(\eta_{n2} > \epsilon \middle| \{\boldsymbol{X}_1 = \boldsymbol{x}\}\right) \\
&\leq E_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\eta_{n2}^2\right)/\epsilon^2 = 1/\left(n_2\epsilon^2\right) = o(1),
\end{aligned}$$

hence, $n_2\eta_{n2}^2\mathrm{I}_{\{\eta_{n2}>\epsilon\}} = o_p(1)$. According to equation (S2.65), we know (S2.64) holds.

Thus, we infer that

$$\left.\frac{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}}{E^{1/2}_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)}\right| \{\boldsymbol{X}_1 = \boldsymbol{x}\} \xrightarrow{d} \mathbf{N}(0,1).$$

The proof is finished. □

The following lemma can be similarly proved by the arguments used in the above proof and is useful in the proof of Lemma 9 and 10.

Lemma 8. *Under Condition 9,*

$$\frac{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}}{E^{1/2}_{\boldsymbol{X}_1}\left(\varepsilon_2^2 w'^2_{12}\right)} \xrightarrow{d} \mathbf{N}(0,1).$$

Lemma 9. *Under Conditions 1-9 and the null hypothesis in (S2.6),*

$$\frac{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2}\xrightarrow{p}1. \tag{S2.66}$$

Proof of Lemma 9. We can prove Lemma 9 through the following four steps:

Step 1. The denominator of equation (S2.66) can be proved that

$$\begin{aligned}&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2\\ =&E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}+o_p(1).\end{aligned} \tag{S2.67}$$

Step 2. Define $\sigma^2_{w,\boldsymbol{X}_1}=E_{\boldsymbol{X}_1}\left(\varepsilon_2^2 w'^2_{12}\right)$. We show that $\frac{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}}{\sigma_{w,\boldsymbol{X}_1}}\xrightarrow{d}\mathbf{N}(0,1)$.

Step 3. We proceed to show that for any $0<\epsilon<\min\{1,C_3\}$, as $n_2\to\infty$,

$$\mathbb{P}\left(E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}\geq(C_3-\epsilon)\,z^2_{\frac{1}{2}+\tilde{\epsilon}}\right)\geq 1-\epsilon, \tag{S2.68}$$

where $\tilde{\epsilon}=\epsilon^3/C_4$ and C_3 and C_4 are two positive constants defined in Condition 9. Here we denote z_α as the α quantile of the standard normal distribution.

Step 4. We finally derive that for any random variable sequence $\xi_n=o_p(1)$, which implies

$$\frac{|\xi_n|}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}=o_p(1).$$

Through *Step 1*, equation (S2.66) can be rewritten as

$$\begin{aligned}&\frac{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2}\\ =&\frac{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}+o_p(1)}.\end{aligned}$$

Step 2-4 ensure that for any random variable sequence $\xi_n = o_p(1)$, it holds that

$$\frac{\xi_n}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}} = o_p(1). \tag{S2.69}$$

All analyses above implies that

$$\begin{aligned}&\frac{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}+o_p(1)}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}\\ =&1+\frac{o_p(1)}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}\xrightarrow{p}1.\end{aligned} \tag{S2.70}$$

According to the Continuous Mapping Theorem, when $X_n \xrightarrow{p} 1$, it holds that $X_n^{-1} \xrightarrow{p} 1$. Thus, (S2.70) implies that

$$\begin{aligned}&\frac{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2}\\ =&\frac{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}+o_p(1)}\xrightarrow{p}1.\end{aligned}$$

The proof of equation (S2.66) is finished.

Step 2 is the result of Lemma 8. Now we discuss *Step 1, Step 3* and *Step 4* one by one.

Prove *Step 1.* Obviously,

$$\begin{aligned}&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2\\ =&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2-\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2\end{aligned}$$

We discuss $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2$ and $\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2$ separately in the following. Recall $R_1^* := R_1 - \tilde{R}_1$. From (S2.60), we know that

$$\begin{aligned}&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2\\ =&\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}+\varepsilon_i R_i^*\right)^2\\ =&\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}\right)^2+\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 R_i^{*2}\\ &+\frac{2}{n_1}\sum_{i=1}^{n_1}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}\right)\varepsilon_i R_i^*. \qquad \text{(S2.71)}\end{aligned}$$

From (S2.62) $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 R_i^{*2} = o_p(1)$. For the last term in equation (S2.71), using Cauchy-Schwarz inequality,

$$\left|\frac{2}{n_1}\sum_{i=1}^{n_1}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}\right)\varepsilon_i R_i^*\right|$$

$$\leq 2\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}^{1/2}\left(\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 R_i^{*2}\right)^{1/2} = o_p(1). \tag{S2.72}$$

The reason for the last equation in (S2.72) is as follows: from Condition 9, the denominator $E\left(\varepsilon_1^2\varepsilon_2^2 w'^2_{12}\right) \leq E^{1/2}\left(\varepsilon_1^4\varepsilon_2^4 w'^4_{12}\right) = \sqrt{C_4}$, and then $E\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\} = E\left(\varepsilon_1^2\varepsilon_2^2 w'^2_{12}\right) \leq \sqrt{C_4}$. Combining with equation (S2.62), equation (S2.72) is proved.

For any $\epsilon > 0$, using Chebyshev inequality, we have

$$\begin{aligned}
&\mathbb{P}\left\{\left|\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}\right)^2 - E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}\right| > \epsilon\right\} \\
\leq &\frac{1}{n_1\epsilon^2}E\left(\left[\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2 - E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}\right]^2\right) \\
\leq &\frac{1}{n_1\epsilon^2}E\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\} = O(1/n_1) = o(1),
\end{aligned}$$

where the two inequalities hold because for any random variable sequence $\boldsymbol{Z_1}$ and $\boldsymbol{Z}_2$,

$E\left\{\boldsymbol{Z}_1 - E_{\boldsymbol{Z}_2}(\boldsymbol{Z}_1)\right\}^2 = E\left[E_{\boldsymbol{Z}_2}\left\{\boldsymbol{Z}_1 - E_{\boldsymbol{Z}_2}(\boldsymbol{Z}_1)\right\}^2\right] \leq E\left\{E_{\boldsymbol{Z}_2}\left(\boldsymbol{Z}_1^2\right)\right\} = E\left(\boldsymbol{Z}_1^2\right)$. So

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}\right)^2 = E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\} + o_p(1).$$

Similarly, $\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = o_p(1)$. See equation (S2.123) in the proof

Lemma 16 for details. Thus,

$$\begin{aligned}&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2\\ =&E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right)^2\right\}+o_p\left(1\right).\end{aligned}$$

Then equation (S2.67) is proved, and so is *Step 1*.

Prove *Step 3*. Recall Condition 9 that assumes $E\left(\varepsilon_1^2\varepsilon_2^2{w'}_{12}^2\right)\geq C_3$, and $E\left(\varepsilon_1^4\varepsilon_2^4{w'}_{12}^4\right)\leq C_4$, where C_3 and C_4 are two positive constants. Remind the definition of $\sigma^2_{w,\boldsymbol{X}_1}$ in *Step 2*, that is, $\sigma^2_{w,\boldsymbol{X}_1}=E_{\boldsymbol{X}_1}\left(\varepsilon_2^2{w'}_{12}^2\right)$. Then $E\left(\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}\right)=E\left(\varepsilon_1^2\varepsilon_2^2{w'}_{12}^2\right)\geq C_3>\epsilon$. According to the uniform convergence in distribution, using the result of *Step 2*, we have

$$\sup_{-\infty<t<\infty}\left|\mathbb{P}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j{w'}_{1j}<t\sigma_{w,\boldsymbol{X}_1}\right\}-\Phi(t)\right|\to 0,\ \text{as}\, n_2\to\infty,$$

where $\Phi(\cdot)$ is the standard normal distribution function. Let z_α denote α quantile of standard normal distribution.

Recall $\tilde{\epsilon}=\frac{\epsilon^3}{2C_4}$. Let $t=z_{\frac{1}{2}+\tilde{\epsilon}}$ and $z_{\frac{1}{2}-\tilde{\epsilon}}$, respectively. To show that $E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right)^2\right\}$ is not "too small", we define a "good" event

$$\mathcal{E}=\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right)^2\geq\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}z^2_{\frac{1}{2}+\tilde{\epsilon}}\right\}.$$

Write $\mathcal{E}^c$ as the complement of $\mathcal{E}$. Now we show when n_2 is large enough,

$\mathbb{P}(\mathcal{E}^c)$ is small. As $n_2 \to \infty$,

$$\begin{aligned}&\mathbb{P}(\mathcal{E}^c) - \left\{\Phi\left(z_{\frac{1}{2}+\tilde{\epsilon}}\right) - \Phi\left(z_{\frac{1}{2}-\tilde{\epsilon}}\right)\right\}\\ =&\mathbb{P}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}\right| < \sigma_{w,\boldsymbol{X}_1} z_{\frac{1}{2}+\tilde{\epsilon}}\right\} - \left\{\Phi\left(z_{\frac{1}{2}+\tilde{\epsilon}}\right) - \Phi\left(z_{\frac{1}{2}-\tilde{\epsilon}}\right)\right\} \to 0.\end{aligned}$$

Note that $\Phi\left(z_{\frac{1}{2}+\tilde{\epsilon}}\right) - \Phi\left(z_{\frac{1}{2}-\tilde{\epsilon}}\right) = 2\tilde{\epsilon}$. That is to say, as $n_2 \to \infty$, $\mathbb{P}(\mathcal{E}^c) \to 2\tilde{\epsilon}$. If the "good" event $\mathcal{E}$ happens, then $\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2 \geq \varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 z_{\frac{1}{2}+\tilde{\epsilon}}^2$. Thus,

$$\begin{aligned}&E_{\mathcal{N}_2}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\\ \geq& E_{\mathcal{N}_2}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\mathrm{I}_{\mathcal{E}}\right)^2 \geq z_{\frac{1}{2}+\tilde{\epsilon}}^2 E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\mathrm{I}_{\mathcal{E}}\right),\end{aligned}$$

and we just need to show that $z_{\frac{1}{2}+\tilde{\epsilon}}^2 E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\mathrm{I}_{\mathcal{E}}\right)$ has a lower bound. Obviously,

$$z_{\frac{1}{2}+\tilde{\epsilon}}^2 E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\mathrm{I}_{\mathcal{E}}\right) = z_{\frac{1}{2}+\tilde{\epsilon}}^2 E\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\right) - z_{\frac{1}{2}+\tilde{\epsilon}}^2 E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\mathrm{I}_{\mathcal{E}^c}\right). \quad \text{(S2.73)}$$

The first term in equation (S2.73) is no less than $C_3 z_{\frac{1}{2}+\tilde{\epsilon}}^2$. Next, we show the second term in equation (S2.73) is small enough. Using Markov inequality, we have

$$\mathbb{P}\left(E_{\mathcal{N}_2}(\mathrm{I}_{\mathcal{E}^c}) \geq \epsilon^2/C_4\right) \leq E(\mathrm{I}_{\mathcal{E}^c})/(\epsilon^2/C_4) \to 2\tilde{\epsilon}/(\epsilon^2/C_4) = \epsilon.$$

That is to say, $\mathbb{P}\left(E_{\mathcal{N}_2}\left(\mathrm{I}_{\mathcal{E}^c}\right) < \epsilon^2/C_4\right) \geq 1-\epsilon$ for n_1, n_2 large enough. Under

Condition 9,

$$E\left(\varepsilon_1^4\sigma_{w,\boldsymbol{X}_1}^4\right)=E\left\{\varepsilon_1^4E_{\boldsymbol{X}_1}^2\left(\varepsilon_2^2{w'}_{12}^2\right)\right\}\leq E\left\{\left(\varepsilon_1\varepsilon_2{w'}_{12}\right)^4\right\}\leq C_4.$$

Then using Cauchy-Schwarz inequality, we have that for large n_1 and n_2, with probability no less than $1-\epsilon$, $E_{\mathcal{N}_2}\left(\mathrm{I}_{\mathcal{E}^c}\right)<\epsilon^2/C_4$ and

$$\begin{aligned}z_{\frac{1}{2}+\tilde{\epsilon}}^2E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\mathrm{I}_{\mathcal{E}^c}\right)\leq&z_{\frac{1}{2}+\tilde{\epsilon}}^2E^{1/2}\left(\varepsilon_1^4\sigma_{w,\boldsymbol{X}_1}^4\right)E_{\mathcal{N}_2}^{1/2}\left(\mathrm{I}_{\mathcal{E}^c}\right)\\ \leq&z_{\frac{1}{2}+\tilde{\epsilon}}^2\sqrt{C_4}\sqrt{\epsilon^2/C_4}=\epsilon z_{\frac{1}{2}+\tilde{\epsilon}}^2.\end{aligned}\tag{S2.74}$$

Incorporating (S2.73) with (S2.74), we derive that as n_1 and n_2 go to infinity, with probability no less than $1-\epsilon$,

$$\begin{aligned}E_{\mathcal{N}_2}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right)^2\geq&z_{\frac{1}{2}+\tilde{\epsilon}}^2E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\mathrm{I}_{\mathcal{E}}\right)\\ \geq&z_{\frac{1}{2}+\tilde{\epsilon}}^2E\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\right)-\epsilon z_{\frac{1}{2}+\tilde{\epsilon}}^2\geq\left(C_3-\epsilon\right)z_{\frac{1}{2}+\tilde{\epsilon}}^2>0.\end{aligned}$$

The proof of (S2.68) is finished.

Prove *Step 4*. For any random variable sequence $\xi_n=o_p\left(1\right)$, we can find a positive constant N_ξ such that for any $n>N_\xi$,

$$\mathbb{P}\left(|\xi_n|\geq\epsilon\left(C_3-\epsilon\right)z_{\frac{1}{2}+\tilde{\epsilon}}^2\right)<\epsilon.$$

Hence, for n and n_2 large enough,

$$\mathbb{P}\left(\frac{|\xi_n|}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right)^2\right\}}>\epsilon\right)$$

$$\leq \mathbb{P}\left(|\xi_n| \geq \epsilon\left(C_3-\epsilon\right) z^2_{\frac{1}{2}+\tilde{\epsilon}}\right) + \mathbb{P}\left(E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j {w'}_{1j}\right)^2\right\} < \left(C_3-\epsilon\right) z^2_{\frac{1}{2}+\tilde{\epsilon}}\right) < 2\epsilon.$$

Then we have that $\frac{|\xi_n|}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j {w'}_{1j}\right)^2\right\}} = o_p\left(1\right)$.

Finally, *Step 1-4* are proved completely, and so is the whole proof. □

Lemma 10 is useful for applying the Berry-Esseen condition to show that our statistic has a standard normal weak limit under the null hypothesis. Lemma 13 is valid for the study under the local alternative hypothesis.

Lemma 10. *Under the null hypothesis in (S2.6) and Conditions 9 holds,*

$$\frac{E_{\mathcal{N}_2}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j {w'}_{1j}\right|^3\right\}}{E^{3/2}_{\mathcal{N}_2}\left\{\log^2 n_1\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j {w'}_{1j}\right)^2\right\}} = O_p\left(1\right). \tag{S2.75}$$

Proof of Lemma 10. We prove Lemma 10 in two steps: first, the denominator in this lemma is no smaller than a constant in probability, and second, the numerator in this lemma is bounded in probability.

Step 1. From Lemma 8 and the uniform convergence in distribution, we have, as $n_2 \to \infty$,

$$\sup_{-\infty<t<\infty}\left|\mathbb{P}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j {w'}_{1j} < t\sigma_{w,\boldsymbol{X}_1}\right\} - \Phi(t)\right| \to 0,$$

where $\Phi(\cdot)$ is the standard normal distribution function. By calculating the

integral of the standard normal distribution density function, we obtain

$$\Phi\left(\frac{1}{\log n_1}\right)-\Phi\left(-\frac{1}{\log n_1}\right)\leq\frac{2}{\sqrt{2\pi}\log n_1}\to 0,\ \text{as}\ n_1\to\infty.$$

Recalling (S2.76), we define the "good" event:

$$\mathcal{A}:=\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right|>\frac{|\varepsilon_1|\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\},$$

and then we start to prove $\mathbb{P}(\mathcal{A})\to 1$. As $n_1\to\infty$ and $n_2\to\infty$,

$$\begin{aligned}
\mathbb{P}(\mathcal{A}^c)=&\mathbb{P}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right|\leq\frac{|\varepsilon_1|\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\}\\
=&\mathbb{P}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j{w'}_{1j}\right|\leq\frac{\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\}\\
=&\mathbb{P}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j{w'}_{1j}\leq\frac{\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\}\\
&-\mathbb{P}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j{w'}_{1j}<-\frac{\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\}\\
\leq&\mathbb{P}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j{w'}_{1j}\leq\frac{\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\}-\Phi\left(\frac{1}{\log n_1}\right)+\Phi\left(-\frac{1}{\log n_1}\right)\\
&-\mathbb{P}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j{w'}_{1j}<-\frac{\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\}+\frac{2}{\sqrt{2\pi}\log n_1}\to 0. \quad \text{(S2.76)}
\end{aligned}$$

From the definition of $\mathcal{A}$, we have

$$\begin{aligned}
&E_{\mathcal{N}_2}\left\{\log^2 n_1\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right)^2\right\}\\
\geq&E_{\mathcal{N}_2}\left\{\log^2 n_1\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j{w'}_{1j}\right)^2\mathrm{I}_{\mathcal{A}}\right\}\\
\geq&E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\mathrm{I}_{\mathcal{A}}\right). \quad \text{(S2.77)}
\end{aligned}$$

Now we show $E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}}\right)$ in (S2.77) has a lower bound away from zero in probability. Note that $E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\right) \geq C_3$ and

$$\begin{aligned} E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}}\right) =& E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\right) - E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}^c}\right) \\ \geq& C_3 - E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}^c}\right). \end{aligned}$$

Then we just need to show $E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}^c}\right) = o_p(1)$. Under Condition 9,

$$\begin{aligned} E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}^c}\right) \leq& E^{1/2}\left(\varepsilon_1^4\sigma_{w,\boldsymbol{X}_1}^4\right) E_{\mathcal{N}_2}^{1/2}\left(\mathrm{I}_{\mathcal{A}^c}\right) \\ \leq& E^{1/2}\left(\varepsilon_1^4\varepsilon_2^4 w'^4_{12}\right) \mathbb{P}^{1/2}\left(\mathcal{A}^c|\mathcal{N}_2\right) \\ \leq& \sqrt{C_4}\mathbb{P}^{1/2}\left(\mathcal{A}^c|\mathcal{N}_2\right), \end{aligned}$$

where the first inequality is derived from Cauchy-Schwarz inequality. According to equation (S2.76), $\mathbb{P}(\mathcal{A}^c) \to 0$ as $n_1 \to \infty$ and $n_2 \to \infty$. For $\mathbb{P}\left(\mathcal{A}^c|\mathcal{N}_2\right)$, which is $\mathcal{N}_2$ measurable, it holds that for any $\epsilon > 0$,

$$\mathbb{P}\{\mathbb{P}\left(\mathcal{A}^c|\mathcal{N}_2\right) > \epsilon\} \leq E\{\mathbb{P}\left(\mathcal{A}^c|\mathcal{N}_2\right)\}/\epsilon = \mathbb{P}\left(\mathcal{A}^c\right)/\epsilon = o(1),$$

then $\mathbb{P}\left(\mathcal{A}^c|\mathcal{N}_2\right) = o_p(1)$, $E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}^c}\right) = o_p(1)$ and $E_{\mathcal{N}_2}\left(\varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 \mathrm{I}_{\mathcal{A}}\right) \geq C_3 + o_p(1)$. From the equation (S2.77), $E_{\mathcal{N}_2}\left\{\log^2 n_1 \left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n} \varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}$ is larger than $C_3/2$ in probability, as n_1 and n_2 goes to infinity.

That is the end of *Step 1*.

Step 2. We now prove that the numerator in (S2.75) is bounded in probability. According to Lyapunov inequality, it is sufficient to prove that

$E_{\mathcal{N}_2}\left\{\left(\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\} = O_p(1)$. By some elementary calculations, we have

$$\begin{aligned}
&E\left\{\left(\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\}\\
=&\sum_{j=n_1+1}^{n}E\left\{\left(\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\}+3\sum_{j_1\neq j_2}E\left\{\left(\varepsilon_1\varepsilon_{j_1} w'_{1j_1}\right)^2\left(\varepsilon_1\varepsilon_{j_2} w'_{1j_2}\right)^2\right\}\\
\leq& n_2E\left\{\left(\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\}+3n_2(n_2-1)E\left\{\left(\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\}.
\end{aligned}$$

The first equation is because all the terms containing $\varepsilon_{j_1}\varepsilon_{j_2}$ or $\varepsilon_{j_1}\varepsilon_{j_2}^3$, $j_1\neq j_2$ in the expectation are all equal to 0.

Further, Condition 9 suggests $E\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\} \leq 3C_4$. According to Markov inequality, we conclude that

$$E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\} = O_p(1).$$

Then it follows that the numerator of (S2.75) is bounded in probability.

That is the end of *Step 2*.

Using the results of *Step 1* and *Step 2*, we conclude that

$$\frac{E_{\mathcal{N}_2}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right|^3\right\}}{E_{\mathcal{N}_2}^{3/2}\left[\log^2 n_1\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right\}^2\right]} = O_p(1).$$

The proof is finished. □

Lemma 11. *Suppose that Conditions 1-5, 7 and 10 hold. Under the global alternative hypothesis in (S2.7) with $\delta_n = 1$, recall $m(\boldsymbol{X}) = E(Y|\boldsymbol{X})$ and*

$l(\boldsymbol{X}) = m(\boldsymbol{X}) - g\left(\boldsymbol{\theta}^*, \boldsymbol{X}\right)$. *If* $p^3 \log n_1/n_1 \to 0$ *and* $p^3 \log n_2/n_2 \to 0$, *then when* $n_1 \to \infty$ *and* $n_2 \to \infty$, *we have that*

$$\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} = E\left\{l(\boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\} + o_p(1).$$

Proof of Lemma 11. Under the global alternative hypothesis,

$$\begin{aligned}
&\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \\
=&\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{\varepsilon_i + m(\boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\} \left\{\varepsilon_j + m(\boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
=&\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{\varepsilon_i + m(\boldsymbol{X}_i) - g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_i\right) + g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_i\right) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\} \\
&\quad \times \left\{\varepsilon_j + m(\boldsymbol{X}_j) - g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) + g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
=&\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w_{ij} + \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_i) l(\boldsymbol{X}_j) w_{ij} \\
&+ \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_i\right) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\} \left\{g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
&+ \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i l(\boldsymbol{X}_j) w_{ij} + \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_i) \varepsilon_j w_{ij} \\
&+ \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \left\{g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
&+ \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_i\right) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\} \varepsilon_j w_{ij} \\
&+ \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_i) \left\{g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} w_{ij} \\
&+ \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_i\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} l(\boldsymbol{X}_j) w_{ij}
\end{aligned}$$

$$=:\mathcal{M}_1+\mathcal{M}_2+\mathcal{M}_3+\mathcal{M}_4+\mathcal{M}_5+\mathcal{M}_6+\mathcal{M}_7+\mathcal{M}_8+\mathcal{M}_9.$$

For $\mathcal{M}_1$,

$$E(\mathcal{M}_1^2)=E\left(\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w_{ij}\right)^2=\frac{1}{n_1n_2}E\left(\varepsilon_1^2\varepsilon_2^2w_{12}^2\right)=O\left(1/(n_1n_2)\right),$$

so $\mathcal{M}_1=o_p(1)$. For $\mathcal{M}_2$,

$$\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_i)l(\boldsymbol{X}_j)w_{ij}=E\left\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}+o_p(1).$$

For $\mathcal{M}_3$,

$$\begin{aligned}\mathcal{M}_3=&\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\left\{g\left(\boldsymbol{\theta}^*,\boldsymbol{X}_j\right)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\\ =&\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left(\boldsymbol{\theta}^*-\hat{\boldsymbol{\theta}}_1\right)^\top\dot{g}\left(\boldsymbol{\theta}^*_{(i)},\boldsymbol{X}_i\right)\left(\boldsymbol{\theta}^*-\hat{\boldsymbol{\theta}}_2\right)^\top\dot{g}\left(\boldsymbol{\theta}^*_{(j)},\boldsymbol{X}_j\right)w_{ij}\\ \le&\left\|\boldsymbol{\theta}^*-\hat{\boldsymbol{\theta}}_1\right\|\left\|\boldsymbol{\theta}^*-\hat{\boldsymbol{\theta}}_2\right\|\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\|\dot{g}\left(\boldsymbol{\theta}^*_{(i)},\boldsymbol{X}_i\right)\right\|\left\|\dot{g}\left(\boldsymbol{\theta}^*_{(j)},\boldsymbol{X}_j\right)\right\|w_{ij}\\ =&O_p\left(\sqrt{p/n_1}\right)O_p\left(\sqrt{p/n_2}\right)O_p(p)=O_p\left(p^2/\sqrt{n_1n_2}\right)=o_p(1),\end{aligned}$$

where $\boldsymbol{\theta}^*_{(i)}$ lies between $\boldsymbol{\theta}^*$ and $\hat{\boldsymbol{\theta}}_1$, and $\boldsymbol{\theta}^*_{(j)}$ lies between $\boldsymbol{\theta}^*$ and $\hat{\boldsymbol{\theta}}_2$. For $\mathcal{M}_4$,

$$\begin{aligned}E\left(\mathcal{M}_4^2\right)=&E\left(\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i l(\boldsymbol{X}_j)w_{ij}\right)^2\\ =&\frac{1}{n_1^2n_2^2}\sum_{i=1}^{n_1}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n}E\left\{\varepsilon_i^2l(\boldsymbol{X}_{j_1})l(\boldsymbol{X}_{j_2})w_{ij_1}w_{ij_2}\right\}\\ \le&\frac{1}{n_1^2n_2^2}\sum_{i=1}^{n_1}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n}E\left\{\varepsilon_i^2l^2(\boldsymbol{X}_{j_1})w_{ij_1}^2\right\}\\ =&\frac{1}{n_1}E\left\{\varepsilon_1^2l^2(\boldsymbol{X}_2)w_{12}^2\right\}\end{aligned}$$

$$\leq \frac{1}{n_1} E^{1/2}\left(\varepsilon_1^4\right) E^{1/2}\left\{l^4(\boldsymbol{X}_2)\right\} E^{1/2}\left(w_{12}^4\right) = O\left(1/n_1\right),$$

where the last inequality is derived from Cauchy-Schwarz inequality. So $\mathcal{M}_4 = o_p(1)$. By similar or easier technique, we can derive that $\mathcal{M}_5 + \mathcal{M}_6 + \mathcal{M}_7 + \mathcal{M}_8 + \mathcal{M}_9 = o_p(1)$. Then

$$\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} = E\left\{l(\boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\} + o_p(1).$$

The proof is finished. □

The following lemma shows the relationship between $\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}$ and $\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij}$ under the local alternative hypothesis.

Lemma 12. *Suppose that Conditions 1-8, 10 and 11 hold. Under the local alternative hypothesis in (S2.7) with* $\delta_n \to 0$, *if* $p^3 \log n_1 / n_1 \to 0$ *and* $p^3 \log n_2 / n_2 \to 0$, *we have the following results.*

(a) If $\delta_n = \frac{1}{\sqrt{n}}$,

$$\begin{aligned}
&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + \frac{\sqrt{n_2}}{\sqrt{n_1 n}} \sum_{i=1}^{n_1} \varepsilon_i H(\boldsymbol{X}_i) \\
&+ \frac{\sqrt{n_1}}{\sqrt{n_2 n}} \sum_{j=n_1+1}^{n} \varepsilon_j H(\boldsymbol{X}_j) + \frac{\sqrt{n_1 n_2}}{n} \mu + o_p(1),
\end{aligned}$$

where

$$\begin{aligned}H(\boldsymbol{X}_1) =& E_{\boldsymbol{X}_1}\left\{l(\boldsymbol{X}_2)w_{12}\right\} - E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)w_{12}\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\} \\ & - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\} \\ & + \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}\end{aligned}$$

and

$$\begin{aligned}\mu =& E\left\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\} \\ & - 2E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\} \\ & + E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}.\end{aligned}$$

(b) If $\delta_n = n^{-\alpha}$, $0 < \alpha < \frac{1}{2}$, $n^{-2\alpha}p^3 \log n \to 0$, *then*

$$\begin{aligned}& \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \\ =& \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + \frac{\sqrt{n_2}}{\sqrt{n_1 n^{2\alpha}}} \sum_{i=1}^{n_1} \varepsilon_i H(\boldsymbol{X}_i) \\ & + \frac{\sqrt{n_1}}{\sqrt{n_2 n^{2\alpha}}} \sum_{j=n_1+1}^{n} \varepsilon_j H(\boldsymbol{X}_j) + \frac{\sqrt{n_1 n_2}}{n^{2\alpha}}\mu \\ & + o_p\left(\max\left\{\sqrt{n_1}n^{-\alpha}, \sqrt{n_2}n^{-\alpha}, \sqrt{n_1 n_2}n^{-2\alpha}\right\}\right).\end{aligned}$$

In particular, when $n_1 = n_2$,

$$\begin{aligned}\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} =& \frac{2}{n} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + \frac{1}{n^\alpha} \sum_{i=1}^{n_1} \varepsilon_i H(\boldsymbol{X}_i) \\ & + \frac{1}{n^\alpha} \sum_{j=n_1+1}^{n} \varepsilon_j H(\boldsymbol{X}_j) + \frac{1}{4} n^{1-2\alpha}\mu + o_p\left(n^{1-2\alpha}\right).\end{aligned}$$

(c) If $\delta_n = n^{-\alpha}$, $\alpha > \frac{1}{2}$, whatever the relationship between n_1 and n_2 as long as they satisfy $p^3 \log n_1/n_1 \to 0$ and $p^3 \log n_2/n_2 \to 0$,

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} = \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + o_p(1).$$

Proof of Lemma 12. To begin with, we decompose

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}$$

as

$$\begin{aligned}
&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{ \varepsilon_i + \delta_n l(\boldsymbol{X}_i) + g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) \right\} \\
&\quad \times \left\{ \varepsilon_j + \delta_n l(\boldsymbol{X}_j) + g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} w_{ij} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w_{ij} + \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n l(\boldsymbol{X}_i) \varepsilon_j w_{ij} \\
&+ \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i l(\boldsymbol{X}_j) w_{ij} + \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n^2 l(\boldsymbol{X}_i) l(\boldsymbol{X}_j) w_{ij} \\
&+ \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} w_{ij} \\
&+ \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) \right\} \varepsilon_j w_{ij} \\
&+ \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n l(\boldsymbol{X}_i) \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} w_{ij} \\
&+ \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) \right\} l(\boldsymbol{X}_j) w_{ij}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\\
&\quad\times\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\\
=:&U_1+\delta_n U_2+\delta_n U_3+\delta_n^2 U_4+U_5+U_6+\delta_n U_7+\delta_n U_8+U_9. \qquad \text{(S2.78)}
\end{aligned}
$$

Next, we discuss U_i, $i=5,6,7,8,9$. For U_5, we notice that

$$
\begin{aligned}
U_5=&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\\
=&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\varepsilon_i\left[\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\right],
\end{aligned}
$$

thus we can use the conclusion of Lemma 3. From Lemma 3 case (c), we have that when $p^3\log n_1/n_1\to 0$, $p^3\log n_2/n_2\to 0$ and $p^3\delta_n^2\log n\to 0$,

$$
\begin{aligned}
&\sum_{j=n_1+1}^{n}\left\{g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)-g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\right\}w_{1j}\\
=&\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\\
&+\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{1j}+\sqrt{n_2}\mathcal{R}_1, \qquad \text{(S2.79)}
\end{aligned}
$$

where $E_{\mathcal{N}_2}\left(\mathcal{R}_1^4\right)=o_p\left(1\right)$. Then it follows that

$$
\begin{aligned}
U_5=&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\varepsilon_i\left[\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{ij}\right]\\
=&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{i2}\right\}\\
&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\delta_n\varepsilon_i\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{ij}
\end{aligned}
$$

$$- \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \mathcal{R}_i, \tag{S2.80}$$

where $\boldsymbol{X}_2$ is a copy of $\boldsymbol{X}$ which is independent of $\boldsymbol{X}_j$ in $E_{\boldsymbol{X}_i}\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{i2}\}$.

By Cauchy-Schwarz inequality,

$$E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \mathcal{R}_i\right)^2\right\} = E_{\mathcal{N}_2}\left(\varepsilon_1^2 \mathcal{R}_1^2\right) \leq E^{1/2}\left(\varepsilon_1^4\right) E_{\mathcal{N}_2}^{1/2}\left(\mathcal{R}_1^4\right) = o_p(1),$$

then according to Markov inequality, $\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \mathcal{R}_i = o_p(1)$. From equation (S2.80) we know that

$$\begin{aligned} U_5 =& - \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_i}\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{i2}\} \\ & - \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\} w_{ij} + o_p(1) \\ =: & \, U_{51} + U_{52} + o_p(1). \end{aligned}$$

Since U_6 has the same structure as U_5,

$$\begin{aligned} U_6 =& - \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\} \\ & - \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\} w_{ij} + o_p(1) \\ =: & \, U_{61} + U_{62} + o_p(1). \end{aligned}$$

For U_7, using (S2.79) again, we have

$$U_7 = \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} l(\boldsymbol{X}_i) \left[\sum_{j=n_1+1}^{n} \left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j)\right\} w_{ij}\right]$$

$$
\begin{aligned}
=&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_j l(\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{i2}\right\}\\
&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}l(\boldsymbol{X}_i)w_{ij}\\
&-\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}l(\boldsymbol{X}_i)\mathcal{R}_i.
\end{aligned}
$$

According to Cauchy-Schwarz inequality,

$$
\begin{aligned}
E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}l(\boldsymbol{X}_i)\mathcal{R}_i\right)^2\right\}\leq&E_{\mathcal{N}_2}\left\{\sum_{i=1}^{n_1}l^2(\boldsymbol{X}_i)\mathcal{R}_i^2\right\}\\
\leq&n_1E^{1/2}\left\{l^4(\boldsymbol{X}_1)\right\}E_{\mathcal{N}_2}^{1/2}\left(\mathcal{R}_1^4\right)=o_p\left(n_1\right),
\end{aligned}
$$

so $\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}l(\boldsymbol{X}_i)\mathcal{R}_i=o_p\left(\sqrt{n_1}\right)$. Thus,

$$
\begin{aligned}
U_7=&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_j l(\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{i2}\right\}\\
&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}l(\boldsymbol{X}_i)w_{ij}\\
&+o_p\left(\sqrt{n_1}\right)\\
=:&\,U_{71}+U_{72}+o_p\left(\sqrt{n_1}\right).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
U_8=&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\right\}\\
&-\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}l(\boldsymbol{X}_j)w_{ij}\\
&+o_p\left(\sqrt{n_2}\right)
\end{aligned}
$$

$$=: U_{81} + U_{82} + o_p\left(\sqrt{n_2}\right).$$

For U_9, according to case (b) in Lemma 4, we know that

$$\begin{aligned}
U_9 =& \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) \right\} \\
& \times \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} w_{ij} \\
=& \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
& + \sqrt{n_1 n_2}\delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
& + \sqrt{n_1 n_2}\delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \\
& + \sqrt{n_1 n_2}\delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
& + o_p\left(\max\left\{1, \sqrt{n_1}\delta_n, \sqrt{n_2}\delta_n, \sqrt{n_1 n_2}\delta_n^2\right\}\right) \\
=: & U_{91} + U_{92} + U_{93} + U_{94} + o_p\left(\max\left\{1, \sqrt{n_1}\delta_n, \sqrt{n_2}\delta_n, \sqrt{n_1 n_2}\delta_n^2\right\}\right).
\end{aligned}$$

Then we have

$$\begin{aligned}
& \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \\
=& U_1 + \delta_n U_2 + \delta_n U_3 + \delta_n^2 U_4 + U_5 + U_6 + \delta_n U_7 + \delta_n U_8 + U_9 \\
=& U_1 + \delta_n U_2 + \delta_n U_3 + \delta_n^2 U_4 + U_{51} + U_{52} + U_{61} + U_{62} + \delta_n U_{71} \\
& + \delta_n U_{72} + \delta_n U_{81} + \delta_n U_{82} + U_{91} + U_{92} + U_{93} + U_{94} \\
& + o_p\left(\max\left\{1, \sqrt{n_1}\delta_n, \sqrt{n_2}\delta_n, \sqrt{n_1 n_2}\delta_n^2\right\}\right). \qquad \text{(S2.81)}
\end{aligned}$$

There are 17 terms in the decomposition of $\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}$, including one remainder term. We combine every 4 of the first 16 terms in the following.

Recall that

$$\begin{aligned}
&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w_{ij} \\
&- \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_i} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{i2} \right\} \\
&- \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j} \right\} \\
&+ \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
=&U_1 + U_{51} + U_{61} + U_{91}. \qquad \text{(S2.82)}
\end{aligned}$$

Thus, the 4 terms in (S2.81) can be written as $\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij}$.

Now we deal with another 4 terms in (S2.81). Here we have the following 3 equations:

$$\begin{aligned}
\delta_n U_3 =&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i l(\boldsymbol{X}_j) w_{ij} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i E_{\boldsymbol{X}_i} \left\{ l(\boldsymbol{X}_2) w_{12} \right\} + o_p(1),
\end{aligned}$$

$$U_{52} = -\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\} w_{ij}$$

$$= -\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i E_{\boldsymbol{X}_i} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{i2} \right\}^\top \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\}$$

$$+ o_p(1). \tag{S2.83}$$

and

$$\begin{aligned}
\delta_n U_{81} =& -\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i l(\boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j} \right\} \\
=& -\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12} \right\} \\
&+ o_p(1).
\end{aligned}$$

We only prove equation (S2.83), and the other two equations can be proved similarly. Recall in the proof of Lemma 3, $\boldsymbol{S}_{1n_2} = \frac{1}{n_2} \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{1j} - E_{\boldsymbol{X}_1} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12} \right\}$ and $E\left(\|\boldsymbol{S}_{1n_2}\|^4 \right) = O\left(p^2/n_2^2\right)$. Thus,

$$\begin{aligned}
&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\} w_{ij} \\
&-\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i E_{\boldsymbol{X}_1} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12} \right\}^\top \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\} \\
=& \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \delta_n \left[\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) w_{ij} - E_{\boldsymbol{X}_i} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{i2} \right\} \right]^\top \boldsymbol{\Sigma}^{-1} \\
&\times E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\} \\
=& \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \sqrt{n_2} \delta_n \boldsymbol{S}_{in_2}^\top \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\}.
\end{aligned}$$

Recall $\left\| \boldsymbol{\Sigma}^{-1} \right\| = O(1)$, and from (S2.39) we know $\| E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\} \| =$

$O(1)$, so we just need to prove that $\left\| \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \sqrt{n_2} \delta_n \boldsymbol{S}_{in_2} \right\| = o_p(1)$. For this term, since $p^3 \delta_n^2 \log n_2 \to 0$, we have

$$
\begin{aligned}
& E \left\| \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \sqrt{n_2} \delta_n \boldsymbol{S}_{in_2} \right\|^2 = n_2 \delta_n^2 E \left(\varepsilon_1^2 \left\| \boldsymbol{S}_{1n_2} \right\|^2 \right) \\
\leq & n_2 \delta_n^2 E^{1/2} \left(\varepsilon_1^4 \right) E^{1/2} \left(\left\| \boldsymbol{S}_{1n_2} \right\|^4 \right) = n_2 \delta_n^2 O\left(p/n_2 \right) = o_p(1).
\end{aligned}
$$

So $\left\| \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \sqrt{n_2} \delta_n \boldsymbol{S}_{in_2} \right\| = o_p(1)$ and (S2.83) holds. Recall $H(\boldsymbol{X}_1) = E_{\boldsymbol{X}_1} \{ l(\boldsymbol{X}_2) w_{12} \} - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12} \}$ and

$$
\begin{aligned}
G(\boldsymbol{X}_1) = & - E_{\boldsymbol{X}_1} \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12} \}^\top \boldsymbol{\Sigma}^{-1} E \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \} \\
& + \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \}.
\end{aligned}
$$

Then we can combine 4 terms in (S2.81) as follows:

$$
\begin{aligned}
& \delta_n U_3 + U_{52} + \delta_n U_{81} + U_{93} \\
= & \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i l(\boldsymbol{X}_j) w_{ij} \\
& - \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \} w_{ij} \\
& - \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_i l(\boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j} \} \\
& + \sqrt{n_1 n_2} \delta_n E \{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \\
= & \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \sqrt{n_2} \delta_n \left\{ H(\boldsymbol{X}_i) + G(\boldsymbol{X}_i) \right\} + o_p\left(\sqrt{n_2} \delta_n \right). \qquad \text{(S2.84)}
\end{aligned}
$$

Similarly, 4 terms in (S2.81) can be written as

$$
\begin{aligned}
&\delta_n U_2 + U_{62} + \delta_n U_{71} + U_{92} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n l(\boldsymbol{X}_i)\varepsilon_j w_{ij} \\
&- \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} w_{ij} \\
&- \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n \varepsilon_j l(\boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{i2}\right\} \\
&+ \sqrt{n_1 n_2}\delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
=&\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \sqrt{n_1}\delta_n \left\{H\left(\boldsymbol{X}_j\right) + G\left(\boldsymbol{X}_j\right)\right\} + o_p\left(\sqrt{n_1}\delta_n\right). \qquad \text{(S2.85)}
\end{aligned}
$$

Finally, we deal with the remaining 4 terms in (S2.81). The discussion for the remainder term is similar to (S2.83):

$$
\begin{aligned}
&\delta_n^2 U_4 + \delta_n U_{72} + \delta_n U_{82} + U_{94} \\
=&\sqrt{n_1 n_2}\delta_n^2 E\left\{l(\boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\} \\
&- 2\sqrt{n_1 n_2}\delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
&+ \sqrt{n_1 n_2}\delta_n^2 E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
&+ o_p\left(\sqrt{n_1 n_2}\delta_n^2\right) \\
=:&\ \sqrt{n_1 n_2}\delta_n^2 \mu + o_p\left(\sqrt{n_1 n_2}\delta_n^2\right). \qquad \text{(S2.86)}
\end{aligned}
$$

Combining (S2.81), (S2.82), (S2.84), (S2.85) and (S2.86), we have

$$\begin{aligned}
&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \sqrt{n_2} \delta_n H(\boldsymbol{X}_i) \\
&+ \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \sqrt{n_1} \delta_n H(\boldsymbol{X}_j) + \sqrt{n_1 n_2} \delta_n^2 \mu \\
&+ o_p\left(\max\left\{1, \sqrt{n_1}\delta_n, \sqrt{n_2}\delta_n, \sqrt{n_1 n_2}\delta_n^2\right\}\right). \qquad \text{(S2.87)}
\end{aligned}$$

Then we use (S2.87) to discuss Conclusions (a) to (c).

About Conclusion (a). If $\delta_n = \frac{1}{\sqrt{n}}$, then we have:

$$\begin{aligned}
&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \\
=&\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + \frac{\sqrt{n_2}}{\sqrt{n_1 n}} \sum_{i=1}^{n_1} \varepsilon_i H(\boldsymbol{X}_i) \\
&+ \frac{\sqrt{n_1}}{\sqrt{n_2 n}} \sum_{j=n_1+1}^{n} \varepsilon_j H(\boldsymbol{X}_j) + \frac{\sqrt{n_1 n_2}}{n} \mu + o_p(1),
\end{aligned}$$

where

$$\begin{aligned}
\mu =&E\left\{l(\boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\} - 2E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
&+ E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\},
\end{aligned}$$

and

$$\begin{aligned}
H(\boldsymbol{X}_1) =&E_{\boldsymbol{X}_1}\left\{l(\boldsymbol{X}_2) w_{12}\right\} - E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12}\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
&- \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\}
\end{aligned}$$

$$+\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}.$$

About Conclusion (b). If $\delta_n = n^{-\alpha}$, $0<\alpha<\frac{1}{2}$, $n^{-2\alpha}p^3\log n\to 0$, then (S2.87) yields that

$$\begin{aligned}
&\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_jw_{ij}\\
=&\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_jw'_{ij}+\frac{\sqrt{n_2}}{\sqrt{n_1n^{2\alpha}}}\sum_{i=1}^{n_1}\varepsilon_iH(\boldsymbol{X}_i)+\frac{\sqrt{n_1}}{\sqrt{n_2n^{2\alpha}}}\sum_{j=n_1+1}^{n}\varepsilon_jH(\boldsymbol{X}_j)\\
&+\frac{\sqrt{n_1n_2}}{n^{2\alpha}}\left\{\mu+o_p\left(\max\left\{\sqrt{n_1}n^{-\alpha},\sqrt{n_2}n^{-\alpha},\sqrt{n_1n_2}n^{-2\alpha}\right\}\right)\right\}.
\end{aligned}$$

Especially when $n_1=n_2$,

$$\begin{aligned}
&\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_jw_{ij}\\
=&\frac{2}{n}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_jw'_{ij}+\frac{1}{n^\alpha}\sum_{i=1}^{n_1}\varepsilon_iH(\boldsymbol{X}_i)\\
&+\frac{1}{n^\alpha}\sum_{j=n_1+1}^{n}\varepsilon_jH(\boldsymbol{X}_j)+\frac{1}{4}n^{1-2\alpha}\mu+o_p\left(\mu+n^{1-2\alpha}\right).
\end{aligned}$$

About Conclusion (c). If $\delta_n = n^{-\alpha}$, $\alpha>\frac{1}{2}$,

$$\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_jw_{ij}=\frac{1}{\sqrt{n_1n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_jw'_{ij}+o_p(1).$$

The proof is finished. □

Lemma 13. *Assume that Conditions 1-11 hold. Under the local alternative hypothesis in (S2.7) with* $\delta_n\to 0$, *if* $p^3\log n_1/n_1\to 0$ *and* $p^3\log n_2/n_2\to 0$,

then

$$\frac{E_{\mathcal{N}_2}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right|^3\right\}}{E_{\mathcal{N}_2}^{3/2}\left[\log^2 n_1\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}=O_p\left(1\right), \tag{S2.88}$$

where

$$\begin{aligned}H(\boldsymbol{X}_1)=&E_{\boldsymbol{X}_1}\left\{l(\boldsymbol{X}_2)w_{12}\right\}-E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}^{\top}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}\\&-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}\\&+\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{*}_{cov}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}.\end{aligned}$$

Proof of Lemma 13. We discuss the case $\sqrt{n_2}\delta_n = O\left(1\right)$ and the case $\sqrt{n_2}\delta_n \to \infty$ separately.

When $\sqrt{n_2}\delta_n = O\left(1\right)$, we prove Lemma 13 in two steps: in probability, the denominator in this lemma is bounded away from zero, and the numerator in this lemma is bounded by a constant.

Step 1. Deal with the denominator. Lemma 7 leads to

$$\left.\frac{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}}{E^{1/2}_{\boldsymbol{X}_1=\boldsymbol{x}}\left(\varepsilon_2^2 w'^2_{12}\right)}\right|\left\{\boldsymbol{X}_1=\boldsymbol{x}\right\}\xrightarrow{d}\mathbf{N}(0,1).$$

Thus,

$$\begin{aligned}&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\Big|\left\{\varepsilon_1,\boldsymbol{X}_1\right\}\\&\xrightarrow{d}\mathbf{N}\left(\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1),\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}\right).\end{aligned}$$

Denote a non-centrality chi-square distribution with the degree of freedom 1 as $\chi_1^2(\delta)$, where δ is the non-centrality parameter. Using the continuous mapping theorem, we have

$$\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\Bigg|\{\varepsilon_1,\boldsymbol{X}_1\}\xrightarrow{d}\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}\chi_1^2\left(H^*(\boldsymbol{X}_1)\right), \tag{S2.89}$$

where $H^*(\boldsymbol{X}_1)=n_2\delta_n^2H^2(\boldsymbol{X}_1)/\sigma^2_{w,\boldsymbol{X}_1}$ is the non-centrality parameter. According to the uniform convergence in distribution, writing the distribution function of $\chi_1^2(\delta)$ as $F_{\chi_1^2(\delta)}(x)$, we have that as $n_2\to\infty$,

$$\begin{aligned}\sup_{-\infty<x<\infty}\Bigg|\mathbb{P}\Bigg\{&\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right)^2\\&<x\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}\Bigg|\{\varepsilon_1,\boldsymbol{X}_1\}\Bigg\}-F_{\chi_1^2(H^*(\boldsymbol{X}_1))}(x)\Bigg|\to 0.\end{aligned} \tag{S2.90}$$

By calculating the integrals, we have that if $n_1\to\infty$,

$$\begin{aligned}F_{\chi_1^2(H^*(\boldsymbol{X}_1))}\left(\frac{1}{\log^2 n_1}\right)\le F_{\chi_1^2(0)}\left(\frac{1}{\log^2 n_1}\right)&=\Phi\left(\frac{1}{\log n_1}\right)-\Phi\left(-\frac{1}{\log n_1}\right)\\&\le\frac{2}{\sqrt{2\pi}\log n_1}\to 0.\end{aligned} \tag{S2.91}$$

Take $x=\frac{1}{\log^2 n_1}$. Then according to (S2.90) and (S2.91), we conclude that

$$\lim_{n_1,n_2\to\infty}\mathbb{P}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right)^2<\frac{\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}}{\log^2 n_1}\Bigg|\{\varepsilon_1,\boldsymbol{X}_1\}\right\}$$

$=0.$

According to the dominated convergence theorem, noting that the probability can be controlled by 1, we have that

$$\begin{aligned}
&\lim_{n_1,n_2\to\infty}\mathbb{P}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right)^2<\frac{\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}}{\log^2 n_1}\right\}=0\\
=&\lim_{n_1,n_2\to\infty}E\left[\mathbb{P}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right)^2\right.\right.\\
&\left.\left.<\frac{\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}}{\log^2 n_1}\middle|\{\varepsilon_1,\boldsymbol{X}_1\}\right\}\right]\\
=&E\left[\lim_{n_1,n_2\to\infty}\mathbb{P}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right)^2\right.\right.\\
&\left.\left.<\frac{\varepsilon_1^2\sigma^2_{w,\boldsymbol{X}_1}}{\log^2 n_1}\middle|\{\varepsilon_1,\boldsymbol{X}_1\}\right\}\right]=0.
\end{aligned}$$

Define an event $\mathcal{B}=\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right|>\frac{|\varepsilon_1|\sigma_{w,\boldsymbol{X}_1}}{\log n_1}\right\}$, which satisfies that $\mathbb{P}(\mathcal{B}^c)\to 0$ as $n_1\to\infty$ and $n_2\to\infty$. Thus, similar to *Step 1* in the proof of Lemma 10, we have that the denominator $E_{\mathcal{N}_2}\left[\log^2 n_1\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]$ is greater than $C_3/2$ in probability, as n_1 and n_2 goes to infinity.

Step 2. Deal with the numerator in (S2.88). According to Lyapunov inequality, we can prove $E_{\mathcal{N}_2}\left[\left\{\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4\right]=O_p(1)$. In the proof of Lemma 10, we have proved that under Condition 9,

$$E_{\mathcal{N}_2}\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^4=O_p(1),$$

and under Condition 11, $E_{\mathcal{N}_2}\left\{|\varepsilon_1 H(\boldsymbol{X}_1)|^4\right\} = O_p(1)$. Hence,

$$E_{\mathcal{N}_2}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4 = O_p(1).$$

Then it follows that the numerator of (S2.88) is bounded in probability.

Using the results of *Step 1*, *Step 2* and Lyapunov inequality, we conclude that

$$\frac{E_{\mathcal{N}_2}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right|^3\right\}}{E_{\mathcal{N}_2}^{3/2}\left[\log^2 n_1\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} = O_p(1).$$

That is the end of the proof of the case $\sqrt{n_2}\delta_n = O(1)$.

When $\sqrt{n_2}\delta_n \to \infty$, since $E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^4\right\} = O_p(1) = o_p(n_2^2\delta_n^4)$, we have

$$\begin{aligned}&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4\right]\\ =&E\left[\left\{\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4\right] + o_p\left(n_2^2\delta_n^4\right),\end{aligned} \tag{S2.92}$$

and

$$\begin{aligned}&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]\\ =&E\left[\left\{\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right] + o_p\left(n_2\delta_n^2\right).\end{aligned} \tag{S2.93}$$

Under Condition 11, we have

$$\frac{E_{\mathcal{N}_2}\left[\left\{\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4\right]}{E_{\mathcal{N}_2}^2\left[\left\{\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} = \frac{E\left[\left\{\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4\right]}{E^2\left[\left\{\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} = O(1), \tag{S2.94}$$

and

$$E\left[\left\{\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right] > n_2\delta_n^2 C_5. \tag{S2.95}$$

Thus, using (S2.92), (S2.93), (S2.94) and (S2.95), we have

$$\begin{aligned}&\frac{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4\right]}{E^2_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}\\ =&\frac{E\left[\left\{\varepsilon_1 H(\boldsymbol{X}_1)\right\}^4\right]+o_p(1)}{E^2\left[\left\{\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]+o_p(1)}=O_p(1).\end{aligned}$$

According to Lyapunov inequality, we derive that

$$\frac{E_{\mathcal{N}_2}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right|^3\right\}}{E^{3/2}_{\mathcal{N}_2}\left[\log^2 n_1\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}=O_p(1).$$

That is the end of the proof of the case $\sqrt{n_2}\delta_n \to \infty$, and so is the whole proof. □

The following lemma extends the result of (S2.72) to the local alternative hypothesis, which is useful in the proof of Lemma 16.

Lemma 14. *Suppose that Conditions 1-11 hold, $p^3\log n_2/n_2 \to 0$. Then we have the following results:*

(a) If $\delta_n = O_p\left(1/\sqrt{n_2}\right)$, then for any random variable sequence $\xi_n = o_p(1)$,

$$\frac{|\xi_n|}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}=o_p(1).$$

(b) If $\sqrt{n_2}\delta_n \to \infty$ *and* $p^3\delta_n^2 \to 0$*, then for any random variable sequence* $\xi_n = o_p\left(n_2\delta_n^2\right)$,

$$\frac{|\xi_n|}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} = o_p(1).$$

Proof of Lemma 14. **Deal with case (a).** The proof is similar to that for Lemma 9. From (S2.89) we have

$$\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2 \Big| \{\varepsilon_1, \boldsymbol{X}_1\} \xrightarrow{d} \varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2\chi_1^2\left(H^*(\boldsymbol{X}_1)\right).$$

For any $\epsilon > 0$, define $\tilde{\epsilon} = \frac{\epsilon^3}{2C_4}$. Write z_α as the α upper quantile of the standard normal distribution. To handle the denominator in case (a), we define an event as

$$\mathcal{K} = \left\{\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2 \geq \varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 z_{\frac{1}{2}+\tilde{\epsilon}}^2\right\}.$$

From (S2.90), we conclude that

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{K}^c \middle| \{\varepsilon_1, \boldsymbol{X}_1\}\right) - F_{\chi_1^2(H^*(\boldsymbol{X}_1))}\left(z_{\frac{1}{2}+\tilde{\epsilon}}^2\right) \\
=&\mathbb{P}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right)^2 < \varepsilon_1^2\sigma_{w,\boldsymbol{X}_1}^2 z_{\frac{1}{2}+\tilde{\epsilon}}^2 \middle| \{\varepsilon_1, \boldsymbol{X}_1\}\right\} \\
&- F_{\chi_1^2(H^*(\boldsymbol{X}_1))}\left(z_{\frac{1}{2}+\tilde{\epsilon}}^2\right) \\
\xrightarrow{p}&0.
\end{aligned}$$

According to the property of non-central chi-square distribution, we have

$$F_{\chi_1^2(H^*(\boldsymbol{X}_1))}\left(z_{\frac{1}{2}+\tilde{\epsilon}}^2\right) \leq F_{\chi_1^2(0)}\left(z_{\frac{1}{2}+\tilde{\epsilon}}^2\right) = \Phi\left(z_{\frac{1}{2}+\tilde{\epsilon}}\right) - \Phi\left(-z_{\frac{1}{2}+\tilde{\epsilon}}\right) = 2\tilde{\epsilon}.$$

Hence, when $n_2 \to \infty$, we have $\mathbb{P}\left(\mathcal{K}^c | \left\{\varepsilon_1, \boldsymbol{X}_1\right\}\right) \leq 2\tilde{\epsilon}$. According to the dominated convergence theorem, since $\mathbb{P}\left(\mathcal{K}^c | \left\{\varepsilon_1, \boldsymbol{X}_1\right\}\right) \leq 1$, we have $\mathbb{P}\left(\mathcal{K}^c\right) = E\left\{\mathbb{P}\left(\mathcal{K}^c | \left\{\varepsilon_1, \boldsymbol{X}_1\right\}\right)\right\} \leq 2\tilde{\epsilon}$ for n_2 large enough. The other discussions are very similar to *Step 3* and *Step 4* in the proof of Lemma 9, and here we omit them.

Case (a) is proved.

Deal with case (b). From (S2.93) and (S2.95), we conclude that case (b) holds. That is the end of our proof. □

Lemma 15. *For random matrices* $\xi_{1n}, \xi_{2n}, \cdots, \xi_{rn}$ *and* $\xi_1, \xi_2, \cdots, \xi_r$, *if* $\|\xi_{1n} - \xi_1\| = o_p(1), \|\xi_{2n} - \xi_2\| = o_p(1), \cdots, \|\xi_{rn} - \xi_r\| = o_p(1)$ *and* $\|\xi_1\| = O_p(1), \|\xi_2\| = O_p(1), \cdots, \|\xi_r\| = O_p(1)$, *then* $\|\xi_{1n}\xi_{2n}\cdots\xi_{rn} - \xi_1\xi_2\cdots\xi_r\| = o_p(1)$.

Proof of Lemma 15. We prove that $r = 2$ as the case that $r \geq 3$ can be similarly treated. Note that

$$\begin{aligned}
\|\xi_{1n}\xi_{2n} - \xi_1\xi_2\| \leq & \|\xi_{1n}\xi_{2n} - \xi_1\xi_{2n}\| + \|\xi_1\xi_{2n} - \xi_1\xi_2\| \\
= & \|(\xi_{1n} - \xi_1)\xi_{2n}\| + \|\xi_1(\xi_{2n} - \xi_2)\| \\
\leq & \|(\xi_{1n} - \xi_1)\xi_2\| + \|(\xi_{1n} - \xi_1)(\xi_{2n} - \xi_2)\| + \|\xi_1(\xi_{2n} - \xi_2)\| \\
= & o_p(1).
\end{aligned}$$

That is the end of our proof. □

The following lemma shows that the sample variances of $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)$ and $\hat{e}_1\tilde{\tilde{w}}_1(\mathcal{N}_2)$ are close to each other sufficiently($\delta_n = 0$ under the null hypothesis).

Lemma 16. *Suppose that Conditions 1-9 hold,* $p^3\log n_1/n_1 \to 0$ *and* $p^3\log n_2/n_2 \to 0$.

(a) Under the null hypothesis in (S2.6),

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}} - 1 = o_p(1).$$

(b) Suppose that Conditions 10 and 11 hold. Under the local alternative hypothesis in (S2.7) with $\delta_n \to 0$, *if* $p^3\delta_n^2\log n \to 0$,

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} - 1 = o_p(1).$$

Proof of Lemma 16. Noting that under the null hypothesis, we can take $\delta_n = 0$ in (S2.7). Thus, we only need to prove case (b) since case (a) is a direct result of case (b). We prove this lemma in 4 steps to respectively deal with the four results:

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\}^2 = E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right] + o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right),$$

$$\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right),$$

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2-\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2=o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right),$$

and

$$\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2-\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2=o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right).$$

From these results, we can then derive that

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2-\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2-\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)\\
=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right),
\end{aligned}$$

which is useful for our conclusion.

Step 1. Prove

$$\begin{aligned}
\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]\\
&+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right).
\end{aligned}$$

Decompose $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2$ as

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 \qquad\qquad (S2.96)\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{\varepsilon_i+g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)+\delta_n l(\boldsymbol{X}_i)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2
\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2+\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2\\
&+\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\tilde{w}_i^2(\mathcal{N}_2)+\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\delta_n l(\boldsymbol{X}_i)\tilde{w}_i^2(\mathcal{N}_2)\\
&+\frac{2}{n_1}\sum_{i=1}^{n_1}\delta_n l(\boldsymbol{X}_i)\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\tilde{w}_i^2(\mathcal{N}_2)\\
&+\frac{1}{n_1}\sum_{i=1}^{n_1}\delta_n^2 l^2(\boldsymbol{X}_i)\tilde{w}_i^2(\mathcal{N}_2)\\
=:&\ D_1+D_2+D_3+D_4+D_5+D_6. \qquad \text{(S2.97)}
\end{aligned}
$$

We now prove

$$
\begin{aligned}
D_1=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]\\
&+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right),
\end{aligned}
$$

and $D_i=o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$, $i=2,3,4,5,6$. We give the proof details of D_i, $i=1,2,6$, and then we can gain the order of D_i, $i=3,4,5$ by Cauchy-Schwarz inequality.

Before the discussion on $D_1=\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2$, we make an analysis about $\tilde{w}_1(\mathcal{N}_2)$. Recall

$$
\begin{aligned}
\tilde{w}_1(\mathcal{N}_2)=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}\right]\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j w_{1j}-\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}\\
&\times\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\}.
\end{aligned}
$$

For $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j w_{1j}$ in $\tilde{w}_1(\mathcal{N}_2)$, similar to (S2.52), according to case (c) in Lemma 3 , we have

$$\begin{aligned}
&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j w_{1j}\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{\varepsilon_j+g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)+\delta_n l(\boldsymbol{X}_j)\right\}w_{1j}\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{1j}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)w_{1j}\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{1j}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)w_{1j}-\mathcal{R}_1,
\end{aligned}\tag{S2.98}$$

where the remainder term $\mathcal{R}_1$ satisfies $E_{\mathcal{N}_2}\left(\mathcal{R}_1^4\right)=o_p\left(1\right)$. Recall $E_{\mathcal{N}_2}\left(\cdot\right)=E\left(\cdot|\mathcal{N}_2\right)$ means the conditional expectation given $\mathcal{N}_2$. Before the discussion about the second term in $\tilde{w}_1(\mathcal{N}_2)$, we define

$$\tilde{\mathcal{R}}_1=\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)^{\top}\boldsymbol{\Sigma}^{-1}E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\mathcal{R}_1\right\}.$$

Then using (S2.98), we have

$$\begin{aligned}
&\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\}\\
=&\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}\right\}
\end{aligned}$$

$$
\begin{aligned}
&- \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\Bigg[\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \\
&\quad \times E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12}\right\}\Bigg] \\
&- \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\Bigg[\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \\
&\quad \times E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} w_{1j}\Bigg] \\
&+ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j) w_{1j}\right\} \\
&- \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) R_1\right\} \\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} \\
&- \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
&- \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \\
&\quad \times E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\
&+ \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} - \tilde{\mathcal{R}}_1,
\end{aligned}
\tag{S2.99}
$$

where $\boldsymbol{\Sigma}^*_{cov} = E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)^\top w_{12}\right\}$. According to (S2.98) and (S2.99), we have

$$\tilde{w}_1(\mathcal{N}_2)$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_2)w_{12}\right\}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{1j}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)w_{1j}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\right\}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{*}_{cov}\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\right\}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}\\
&\quad\times E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\right\}-\mathcal{R}_1+\tilde{\mathcal{R}}_1\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)w_{1j}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}w_{1j}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\right\}\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_j)^{\top}\boldsymbol{\Sigma}^{-1}\\
&\quad\times E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)^{\top}\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)w_{1j}\right\}+\mathcal{R}_1^{*},
\end{aligned}
$$

where $\mathcal{R}_1^{*}=-\mathcal{R}_1+\tilde{\mathcal{R}}_1$. Similarly, we define $\mathcal{R}_i^{*}$ for $i\neq 1$. To simplify the

other terms in $\tilde{w}_1(\mathcal{N}_2)$, we use the following four notations:

$$Q_1^{(1)} := \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)w_{1j} - E_{\boldsymbol{X}_1}\left\{l(\boldsymbol{X}_2)w_{12}\right\},$$

$$\begin{aligned} Q_1^{(2)} :=& \frac{1}{n_2}\sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} w_{1j} \\ & - E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12}\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}, \end{aligned}$$

$$\begin{aligned} Q_1^{(3)} :=& \frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} \\ & - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\}, \end{aligned}$$

$$\begin{aligned} Q_1^{(4)} :=& \frac{1}{n_2}\sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \quad \text{(S2.100)} \\ & \times \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\ & - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}. \quad \text{(S2.101)} \end{aligned}$$

Let $Q_1 = Q_1^{(1)} + Q_1^{(2)} + Q_1^{(3)} + Q_1^{(4)}$. Similarly, we can define Q_i for $i \neq 1$.

Recall

$$\begin{aligned} H(\boldsymbol{X}_1) =& E_{\boldsymbol{X}_1}\left\{l(\boldsymbol{X}_2)w_{12}\right\} - E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2) w_{12}\right\}^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} \\ & - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12}\right\} \\ & + \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}. \end{aligned}$$

Thus, we have

$$\tilde{w}_1(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n} \varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n H(\boldsymbol{X}_1) + \mathcal{R}_1^* + \sqrt{n_2}\delta_n Q_1.$$

First, analyse D_1. We have

$$\begin{aligned}
D_1 =& \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i \tilde{w}_i(\mathcal{N}_2)\right\}^2 \\
=& \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + \sqrt{n_2}\delta_n\varepsilon_i H(\boldsymbol{X}_i) + \varepsilon_i\left(\mathcal{R}_i^* + Q_i\right)\right\}^2 \\
=& \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + \sqrt{n_2}\delta_n\varepsilon_i H(\boldsymbol{X}_i)\right\}^2 \\
&+ \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\left(\mathcal{R}_i^* + \sqrt{n_2}\delta_n Q_i\right)\right\}^2 \\
&+ \frac{2}{n_1}\sum_{i=1}^{n_1}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + \sqrt{n_2}\delta_n\varepsilon_i H(\boldsymbol{X}_i)\right\}\varepsilon_i\left(\mathcal{R}_i^* + \sqrt{n_2}\delta_n Q_i\right),
\end{aligned} \tag{S2.102}$$

The first term in (S2.102) can be derived as

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + \sqrt{n_2}\delta_n\varepsilon_i H(\boldsymbol{X}_i)\right\}^2 \\
=& E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right] + o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right) \\
=& O_p\left(\max\left\{1, n_2\delta_n^2\right\}\right),
\end{aligned} \tag{S2.103}$$

where lines 2 and 3 are because under Conditions 9 and 11, we have

$$\begin{aligned}
&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right] \\
\leq& 2E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right\}^2\right] + n_2\delta_n^2 E\left\{\varepsilon_1^2 H^2(\boldsymbol{X}_1)\right\} \\
=& O_p\left(\max\left\{1, n_2\delta_n^2\right\}\right).
\end{aligned}$$

For the second term in (S2.102), recall $\mathcal{R}_i^* = -\mathcal{R}_i + \tilde{\mathcal{R}}_i$. From (S2.34) and Markov inequality, we know that $\left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \right\| = O_p(1)$. Then it follows that

$$
\begin{aligned}
&\frac{1}{n_1} \sum_{i=1}^{n_1} \left(\varepsilon_i \tilde{\mathcal{R}}_i \right)^2 \\
=&\frac{1}{n_1} \sum_{i=1}^{n_1} \left[\varepsilon_i \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \mathcal{R}_1 \right\} \right]^2 \\
=&E_{\mathcal{N}_2} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \mathcal{R}_1 \right\}^\top \boldsymbol{\Sigma}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i) \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \\
&\quad \times \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) \mathcal{R}_1 \right\} \\
=&o_p(1). \qquad \text{(S2.104)}
\end{aligned}
$$

Using Cauchy-Schwarz inequality, we have

$$
E_{\mathcal{N}_2} \left\{ \frac{1}{n_1} \sum_{i=1}^{n_1} \left(\varepsilon_i \mathcal{R}_i \right)^2 \right\} = E_{\mathcal{N}_2} \left\{ \left(\varepsilon_1 \mathcal{R}_1 \right)^2 \right\} \leq E^{1/2} \left(\varepsilon_1^4 \right) E_{\mathcal{N}_2}^{1/2} \left(\mathcal{R}_1^4 \right) = o_p(1).
$$

Then we infer that $\frac{1}{n_1} \sum_{i=1}^{n_1} \left(\varepsilon_i \mathcal{R}_1 \right)^2 = o_p(1)$. So $\frac{1}{n_1} \sum_{i=1}^{n_1} \left(\varepsilon_i \mathcal{R}_i^* \right)^2 = o_p(1)$.

Now we show that $\frac{1}{n_1} \sum_{i=1}^{n_1} \left(\varepsilon_i Q_i^{(k)} \right)^2 = o_p(1)$, $k = 1, 2, 3, 4$. We prove the case $k = 4$ since the cases $k = 1, 2, 3$ can be proved by easier methods. Recall the definition of $Q_i^{(4)}$ in (S2.100). For the convenience, we define $\boldsymbol{M}_{n_2} = \frac{1}{n_2} \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_j} \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j} \right\} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top - \boldsymbol{\Sigma}_{cov}^*$. Thus

$$
Q_1^{(4)} = \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{M}_{n_2} \boldsymbol{\Sigma}^{-1} E \left\{ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1) \right\}.
$$

Then we have

$$\begin{aligned}&\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\varepsilon_i Q_i^{(4)}\right)^2\\ =&E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_{n_2}\boldsymbol{\Sigma}^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^\top\\ &\times\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_{n_2}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}.\end{aligned}$$

From (S2.34), we know $\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^\top\right\| = O_p(1)$. According to (S2.39),

$\|E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}\| = O(1)$. Under Condition 4, we have $\left\|\boldsymbol{\Sigma}^{-1}\right\| = O(1)$. Similar to the proof of (S2.29), we have

$$\|\boldsymbol{M}_{n_2}\| = o_p(1/\sqrt{p}) = o_p(1). \tag{S2.105}$$

Thus, $\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\varepsilon_i Q_i^{(4)}\right)^2 = o_p(1)$.

Since $Q_i = Q_i^{(1)} + Q_i^{(2)} + Q_i^{(3)} + Q_i^{(4)}$ and $\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\varepsilon_i Q_i^{(k)}\right)^2 = o_p(1)$, $k = 1,2,3,4$, we conclude that

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\left(\varepsilon_i Q_i\right)^2 = o_p(1). \tag{S2.106}$$

Then it follows that

$$\begin{aligned}\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\left(\mathcal{R}_i^* + \sqrt{n_2}\delta_n Q_i\right)\right\}^2 &\le \frac{2}{n_1}\sum_{i=1}^{n_1}\left(\varepsilon_i R_i^*\right)^2 + \frac{2}{n_1}\sum_{i=1}^{n_1}\left(\varepsilon_i\sqrt{n_2}\delta_n Q_i\right)^2\\ &= o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right).\end{aligned} \tag{S2.107}$$

For the last term in (S2.102), we have

$$\begin{aligned}
&\frac{2}{n_1}\sum_{i=1}^{n_1}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}+\sqrt{n_2}\delta_n\varepsilon_i H(\boldsymbol{X}_i)\right\}\varepsilon_i\left(\mathcal{R}_i^*+Q_i\right)\\
\leq& 2\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}+\sqrt{n_2}\delta_n\varepsilon_i H(\boldsymbol{X}_i)\right\}^2\right]^{1/2}\\
&\quad\times\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\left(\mathcal{R}_i^*+Q_i\right)\right\}^2\right]^{1/2}\\
=& o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right). \qquad \text{(S2.108)}
\end{aligned}$$

Combining (S2.102), (S2.103), (S2.107) and (S2.108), we have

$$\begin{aligned}
D_1=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)\\
=&O_p\left(\max\left\{1,n_2\delta_n^2\right\}\right). \qquad \text{(S2.109)}
\end{aligned}$$

Second, we proceed to prove $D_2 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. we have the following decomposition:

$$\begin{aligned}
D_2=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^\top\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\}\tilde{w}_i(\mathcal{N}_2)\right.\\
&\left.-\left\{\frac{1}{2}\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^\top\ddot{g}\left(\tilde{\boldsymbol{\theta}}^*_{(i)},\boldsymbol{X}_i\right)\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2\\
\leq&\frac{2}{n_1}\sum_{i=1}^{n_1}\left[\left\{\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^\top\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2\\
&+\frac{2}{n_1}\sum_{i=1}^{n_1}\left[\left\{\frac{1}{2}\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)^\top\ddot{g}\left(\tilde{\boldsymbol{\theta}}^*_{(i)},\boldsymbol{X}_i\right)\left(\hat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_0\right)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2
\end{aligned}$$

$$=: D_{21} + D_{22},$$

where $\boldsymbol{\theta}^*_{(i)}$ lies between $\hat{\boldsymbol{\theta}}_1$ and $\boldsymbol{\theta}_0$.

Now we prove $D_{21} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. Recall

$$\tilde{w}_1(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n H(\boldsymbol{X}_1) + \mathcal{R}_1^* + \sqrt{n_2}\delta_n Q_1,$$

where $\mathcal{R}_1^* = -\mathcal{R}_1 + \tilde{\mathcal{R}}_1$. To deal with the remainder terms $\mathcal{R}_1^*$ and $\sqrt{n_2}\delta_n Q_1$ in $\tilde{w}_1(\mathcal{N}_2)$, we define $\mathcal{T}_1 = \left\{E_{\mathcal{N}_2}\left(\mathcal{R}_1^4\right) \leq 1\right\}$ and $\mathcal{T}_2 = \left\{E_{\mathcal{N}_2}\left(Q_1^4\right) \leq 1\right\}$. From Lemma 3, $\mathbb{P}\left(\mathcal{T}_1^c\right) = o\left(1\right)$. If we can prove that $E\left(\mathcal{R}_1^{*4}\mathrm{I}_{\mathcal{T}_1}\right) = O\left(p^2\right)$ and $\mathbb{P}\left(\mathcal{T}_2^c\right) = o\left(1\right)$, $\mathbb{P}\left(\mathcal{T}_1 \cap \mathcal{T}_2\right) \to 1$. We will prove the above in the proofs of (S2.118) and (S2.119) later. Thus, if we can show $D_{21}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2} = o_p\left(1\right)$, then $D_{21} = o_p\left(1\right)$. According to Markov inequality, we just need to prove $E_{\mathcal{N}_1}\left(D_{21}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right) = o_p\left(1\right)$. Note that

$$\begin{aligned}
&E_{\mathcal{N}_1}(D_{21}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}) \\
=&\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top \frac{2}{n_1}\sum_{i=1}^{n_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right) \\
\leq&\left\|\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right\|^2 \left\|\frac{2}{n_1}\sum_{i=1}^{n_1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}\right\| \\
\leq&2\left\|\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right\|^2 \frac{1}{n_1}\sum_{i=1}^{n_1} \left\|\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\right\|^2 E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}. \qquad \text{(S2.110)}
\end{aligned}$$

From Lemma 2, we know $\left\|\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right\|^2 = O_p\left(\max\left\{p/n_1, p\delta_n^2\right\}\right)$. To get the

order of another term in (S2.110), we compute its expectation as follows:

$$
\begin{aligned}
& E\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)\right\|^2 E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}\right] \\
=& E\left[\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^2\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n H(\boldsymbol{X}_1)+\mathcal{R}_1^*+\sqrt{n_2}\delta_n Q_1\right\}^2\right. \\
& \left.\times\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right] \\
\leq& 8E\left[\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^2\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n H(\boldsymbol{X}_1)\right\}^2\right] \\
& +8E\left\{\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^2\mathcal{R}_1^{*2}\mathrm{I}_{\mathcal{T}_1}\right\}+8E\left\{\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^2\left(\sqrt{n_2}\delta_n Q_1\right)^2\mathrm{I}_{\mathcal{T}_2}\right\}.
\end{aligned}
\tag{S2.111}
$$

For the first term in (S2.111), using Cauchy-Schwarz inequality, we conclude that

$$
\begin{aligned}
& E\left[\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^2\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n H(\boldsymbol{X}_1)\right\}^2\right] \\
\leq& E^{1/2}\left\{\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^4\right\}E^{1/2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n H(\boldsymbol{X}_1)\right\}^4\right] \\
=& O(p)\,O\left(\max\left\{1,n_2\delta_n^2\right\}\right)=O\left(\max\left\{p,pn_2\delta_n^2\right\}\right).
\end{aligned}
\tag{S2.112}
$$

For the second term in (S2.111),

$$
\begin{aligned}
& E\left\{\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^2\mathcal{R}_1^{*2}\mathrm{I}_{\mathcal{T}_1}\right\} \\
\leq& E^{1/2}\left\{\left\|\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)\right\|^4\right\}E^{1/2}\left\{\mathcal{R}_1^{*4}\mathrm{I}_{\mathcal{T}_1}\right\}=O(p)\,O(p)=O\left(p^2\right),
\end{aligned}
\tag{S2.113}
$$

where $E\left(\mathcal{R}_1^{*4}\mathrm{I}_{\mathcal{T}_1}\right)=O\left(p^2\right)$ is derived from (S2.118) as explained above. For

the third term in (S2.111),

$$\begin{aligned} & E\left\{\left\|\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\right\|^2 \left(\sqrt{n_2}\delta_n Q_1\right)^2 \mathrm{I}_{\mathcal{T}_2}\right\} \\ \leq & E^{1/2}\left\{\left\|\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\right\|^4\right\} E^{1/2}\left\{\left(\sqrt{n_2}\delta_n Q_1\right)^4 \mathrm{I}_{\mathcal{T}_2}\right\} \\ = & O(p)\, O\left(n_2\delta_n^2\right) = O\left(pn_2\delta_n^2\right), \end{aligned} \tag{S2.114}$$

where $E\left\{\left(\sqrt{n_2}\delta_n Q_1\right)^4 \mathrm{I}_{\mathcal{T}_2}\right\} = O\left(n_2^2\delta_n^4\right)$ is derived from the definition of $\mathcal{T}_2$.

From (S2.111), (S2.112), (S2.113) and (S2.114), we have

$$E\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\|\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\right\|^2 E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}\right] = O\left(\max\left\{p^2, pn_2\delta_n^2\right\}\right). \tag{S2.115}$$

Thus $E_{\mathcal{N}_1}(D_{21}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}) = O_p\left(\max\left\{p^2, pn_2\delta_n^2\right\}\right)$. From (S2.110), since $p^3\log n_1/n_1 \to 0$, $p^3\log n_2/n_2 \to 0$ and $p^3\delta_n^2\log n \to 0$, we derive that

$$\begin{aligned} E_{\mathcal{N}_1}(D_{21}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}) = & O_p\left(\max\left\{p/n_1, p\delta_n^2\right\}\right) O_p\left(\max\left\{p^2, pn_2\delta_n^2\right\}\right) \\ = & o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right). \end{aligned}$$

Then it follows that $D_{21}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$ and $D_{21} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$.

As commented above, we now prove $E\left(\mathcal{R}_1^{*4}\mathrm{I}_{\mathcal{T}_1}\right) = O\left(p^2\right)$ and $\mathbb{P}\left(\mathcal{T}_2^c\right) = o(1)$ in the following.

From the definition of $\mathcal{T}_1$ we know that $E\left(\mathcal{R}_1^4\mathrm{I}_{\mathcal{T}_1}\right) = E\left\{E_{\mathcal{N}_2}\left(\mathcal{R}_1^4\right)\mathrm{I}_{\mathcal{T}_1}\right\} =$

$o(1)$. As for $E\left(\tilde{\mathcal{R}}_1^4 \mathrm{I}_{\mathcal{T}_1}\right)$, we have

$$\begin{aligned} E\left(\tilde{\mathcal{R}}_1^4 \mathrm{I}_{\mathcal{T}_1}\right) =& E\left(\left[\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right)^{\top} \boldsymbol{\Sigma}^{-1} E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\mathcal{R}_1\right\} \mathrm{I}_{\mathcal{T}_1}\right]^4\right) \\ \leq & E\left(\left\|\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right)\right\|^4 \left\|E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\mathcal{R}_1\right\} \mathrm{I}_{\mathcal{T}_1}\right\|^4\right)\left\|\boldsymbol{\Sigma}^{-1}\right\|^4. \end{aligned} \tag{S2.116}$$

According to Condition 4, $\|\boldsymbol{\Sigma}\| \leq b$. Similar to (S2.57), we have

$$\left\|E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\mathcal{R}_1\right\} \mathrm{I}_{\mathcal{T}_1}\right\|^2 \leq \|\boldsymbol{\Sigma}\| E_{\mathcal{N}_2}\left(\mathcal{R}_1^2\right) \mathrm{I}_{\mathcal{T}_1} \leq b.$$

So $\left\|E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\mathcal{R}_1\right\} \mathrm{I}_{\mathcal{T}_1}\right\|^4 \leq b^2$. From (S2.116) we know

$$E\left(\tilde{\mathcal{R}}_1^4 \mathrm{I}_{\mathcal{T}_1}\right) \leq b^2 E\left\{\left\|\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right)\right\|^4\right\}\left\|\boldsymbol{\Sigma}^{-1}\right\|^4 = O\left(p^2\right). \tag{S2.117}$$

Recall $\mathcal{R}_1^* = -\mathcal{R}_1 + \tilde{\mathcal{R}}_1$. Then

$$E\left(\mathcal{R}_1^{*4} \mathrm{I}_{\mathcal{T}_1}\right) \leq 16E\left(\mathcal{R}_1^4 \mathrm{I}_{\mathcal{T}_1}\right) + 16E\left(\tilde{\mathcal{R}}_1^4 \mathrm{I}_{\mathcal{T}_1}\right) = O\left(p^2\right). \tag{S2.118}$$

Next, we prove $\mathbb{P}\left(\mathcal{T}_2^c\right) = o(1)$. By the definition of the set $\mathcal{T}_2^c$, it is sufficient to prove $E_{\mathcal{N}_2}\left(Q_1^4\right) = o_p(1)$. Recall $Q_1 = Q_1^{(1)} + Q_1^{(2)} + Q_1^{(3)} + Q_1^{(4)}$. Here we just prove $E_{\mathcal{N}_2}\left\{\left(Q_1^{(4)}\right)^4 \mathrm{I}_{\mathcal{T}_2}\right\} = o_p(1)$ as an example since we can prove more easily $E_{\mathcal{N}_2}\left\{\left(Q_1^{(k)}\right)^4 \mathrm{I}_{\mathcal{T}_2}\right\} = o_p(1)$, $k = 1, 2, 3$. Recall

$$Q_1^{(4)} = \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{M}_{n_2} \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\},$$

where $\boldsymbol{M}_{n_2} = \frac{1}{n_2} \sum_{j=n_1+1}^{n} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) w_{1j}\right\} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^{\top} - \boldsymbol{\Sigma}_{cov}^*$. Note that $\boldsymbol{M}_{n_2}$ is $\mathcal{N}_2$ measurable. From (S2.105) we know $\|\boldsymbol{M}_{n_2}\| = o_p\left(1/\sqrt{p}\right)$. Then

we infer that

$$E_{\mathcal{N}_2}\left\{\left(Q_1^{(4)}\right)^4 \mathrm{I}_{\mathcal{T}_2}\right\} \leq E\left\{\|\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\|^4\right\} \left\|\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_{n_2}\boldsymbol{\Sigma}^{-1}E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}\right\|^4$$

$$= O\left(p^2\right) o_p\left(1/p^2\right) = o_p\left(1\right).$$

From the results that $E_{\mathcal{N}_2}\left\{\left(Q_1^{(k)}\right)^4 \mathrm{I}_{\mathcal{T}_2}\right\} = o_p\left(1\right)$, $k = 1, 2, 3, 4$ we conclude that $E_{\mathcal{N}_2}\left(Q_1^4\right) = o_p\left(1\right)$. So

$$\mathbb{P}\left(\mathcal{T}_2^c\right) = o\left(1\right). \tag{S2.119}$$

That is the end of the proof about $D_{21} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$.

Next, prove $D_{22} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. Similar to the discussion of D_{21}, we have

$$E_{\mathcal{N}_1}(D_{22}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}) \tag{S2.120}$$

$$= \frac{2}{n_1}\sum_{i=1}^{n_1}\left\{\frac{1}{2}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^\top \ddot{g}\left(\tilde{\boldsymbol{\theta}}_{(i)}^*, \boldsymbol{X}_i\right)\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)\right\}^2 E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}$$

$$\leq \frac{1}{2n_1}\left\|\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right\|^4 \sum_{i=1}^{n_1}\left\|\ddot{g}\left(\tilde{\boldsymbol{\theta}}_{(i)}^*, \boldsymbol{X}_i\right)\right\|^2 E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}. \tag{S2.121}$$

Lemma 2 implies $\left\|\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right\|^4 = O_p\left(\max\left\{p^2/n_1^2, p^2\delta_n^4\right\}\right)$. Similar to the proof of (S2.115), we have

$$E\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\|\ddot{g}\left(\tilde{\boldsymbol{\theta}}_{(i)}^*, \boldsymbol{X}_i\right)\right\|^2 E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}\right] = O\left(\max\left\{p^3, p^2n_2\delta_n^2\right\}\right).$$

Thus, from (S2.120) we derive that

$$E_{\mathcal{N}_1}(D_{22}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}) = O\left(\max\left\{p^2/n_1^2, p^2\delta_n^4\right\}\right) O\left(\max\left\{p^3, p^2n_2\delta_n^2\right\}\right)$$

$$=O\left(\max\left\{p^5/n_1^2, p^5\delta_n^4, n_2\delta_n^2 p^4/n_1^2, n_2\delta_n^2 p^4\delta_n^4\right\}\right).$$

If $p^3\log n_1/n_1 \to 0$, $p^3\log n_2/n_2 \to 0$ and $p^3\delta_n^2\log n \to 0$, then $E_{\mathcal{N}_1}(D_{22}\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}) = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. So $D_{22} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$.

Thus, $D_2 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$.

Third, we analyse D_6. To prove $D_6 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$, we compute the expectation of $D_6\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}$ first. Similar to the proof of (S2.115), we have

$$\begin{aligned}
E\left(D_6\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right) =& E\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\delta_n^2 l^2(\boldsymbol{X}_i)\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}\\
=& \delta_n^2 E\left[\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)E_{\boldsymbol{X}_i}\left\{\tilde{w}_i^2(\mathcal{N}_2)\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2}\right\}\right]\\
=& O\left(\max\left\{p\delta_n^2, n_2\delta_n^4\right\}\right) = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right).
\end{aligned}$$

Then Markov inequality implies that $D_6\mathrm{I}_{\mathcal{T}_1\cap\mathcal{T}_2} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. Since $\mathbb{P}\left(\mathcal{T}_1\cap\mathcal{T}_2\right)$ goes to 1, we have $D_6 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$.

Further, using Cauchy-Schwarz inequality, we derive that when $p^3\log n_1/n_1 \to 0$, $p^3\log n_2/n_2 \to 0$ and $p^3\delta_n^2\log n \to 0$, $D_3 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$, $D_4 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$ and $D_5 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$.

In conclusion,

$$\begin{aligned}
\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 =& E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]\\
&+ o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right).
\end{aligned}$$

The proof of *Step 1* is finished.

Step 2. Prove that $\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = O_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. We have

$$\begin{aligned}
&\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 \\
=&\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i + g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) + \delta_n l(\boldsymbol{X}_i)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2 \\
\leq& 8\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 + 8\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2 \\
&+ 8\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\delta_n l(\boldsymbol{X}_i)\tilde{w}_i(\mathcal{N}_2)\right\}^2. \qquad \text{(S2.122)}
\end{aligned}$$

For the first term in (S2.122),

$$E_{\mathcal{N}_2}\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = \frac{1}{n_1^2}\sum_{i=1}^{n_1}E_{\mathcal{N}_2}\left\{\varepsilon_i^2\tilde{w}_i^2(\mathcal{N}_2)\right\}. \qquad \text{(S2.123)}$$

Similar to (S2.71) and (S2.72), we have

$$\begin{aligned}
&E_{\mathcal{N}_2}\left\{\varepsilon_1^2\tilde{w}_1^2(\mathcal{N}_2)\right\} \\
=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1) + \varepsilon_1\mathcal{R}_1^* + \sqrt{n_2}\delta_n Q_1\right\}^2\right] \\
=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right] \\
&+ E_{\mathcal{N}_2}\left\{\varepsilon_1\left(\mathcal{R}_1^* + \sqrt{n_2}\delta_n Q_1\right)^2\right\} \\
&+ 2E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}\varepsilon_1\left(\mathcal{R}_1^* + \sqrt{n_2}\delta_n Q_1\right)\right]
\end{aligned}$$

$$=E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$$

$$=O_p\left(\max\left\{1,n_2\delta_n^2\right\}\right), \tag{S2.124}$$

where $\mathcal{R}_1^*$ and Q_1 are defined in *Step 1* and the proofs for $E_{\mathcal{N}_2}\left(\varepsilon_1^2\mathcal{R}_1^{*2}\right)=o_p(1)$ and $E_{\mathcal{N}_2}\left(\varepsilon_1^2 {Q_1}^2\right)=o_p(1)$ are similar to (S2.104) and (S2.106). Then according to (S2.123) and (S2.124), we have

$$E_{\mathcal{N}_2}\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2=O_p\left(\max\left\{1/n_1,n_2\delta_n^2/n_1\right\}\right)=o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right).$$

For the second term in (S2.122), according to Cauchy-Schwarz inequality,

$$\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\tilde{w}_i(\mathcal{N}_2)\right]^2\le D_2=o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right),$$

where D_2 is defined in *Step 1*. For the third term in (S2.122), similar to (S2.124), we can conclude that

$$\begin{aligned}E_{\mathcal{N}_2}\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\tilde{w}_i^2(\mathcal{N}_2)\right\}=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n H(\boldsymbol{X}_1)\right\}^2\right]\\&+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)\\=&O_p\left(\max\left\{1,n_2\delta_n^2\right\}\right).\end{aligned}$$

Then Markov inequality implies that $\frac{1}{n_1}\sum_{i=1}^{n_1}\tilde{w}_i^2(\mathcal{N}_2)=O_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$. Thus,

$$\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\delta_n l(\boldsymbol{X}_i)\tilde{w}_i(\mathcal{N}_2)\right\}^2\le\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\delta_n^2 l^2\left(\boldsymbol{X}_i\right)\right\}\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\tilde{w}_i^2(\mathcal{N}_2)\right\}$$

$$=O_p\left(\max\left\{\delta_n^2, n_2\delta_n^4\right\}\right).$$

Hence, we can derive $\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = O_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$ from (S2.122). The proof of *Step 2* is finished.

Step 3. Prove $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2 - \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. We have already decomposed $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2$ as the sum of D_i, $i = 1,2,3,4,5,6$ in (S2.96). Here, we also decompose $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2$ as the sum of D'_i, $i = 1,2,3,4,5,6$, and we now prove that $D'_i - D_i = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$ in the following. For $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2$, we have

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2 \\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{\varepsilon_i + g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) + \delta_n l(\boldsymbol{X}_i)\right\}\tilde{\tilde{w}}_i(\mathcal{N}_2)\right]^2 \\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2 + \frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\tilde{\tilde{w}}_i(\mathcal{N}_2)\right]^2 \\
&+ \frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\tilde{\tilde{w}}_i^2(\mathcal{N}_2) + \frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\delta_n l(\boldsymbol{X}_i)\tilde{\tilde{w}}_i^2(\mathcal{N}_2) \\
&+ \frac{2}{n_1}\sum_{i=1}^{n_1}\delta_n l(\boldsymbol{X}_i)\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\tilde{\tilde{w}}_i^2(\mathcal{N}_2) \\
&+ \frac{1}{n_1}\sum_{i=1}^{n_1}\delta_n^2 l^2(\boldsymbol{X}_i)\tilde{\tilde{w}}_i^2(\mathcal{N}_2) \\
=:& D'_1 + D'_2 + D'_3 + D'_4 + D'_5 + D'_6.
\end{aligned}$$

Recall

$$\tilde{\tilde{w}}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[w_{ij} - \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right],$$

where $\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top$. We just prove $D'_1 - D_1 = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$ as $D'_i - D_i = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$, $i = 2, 3, 4, 5, 6$ can be proved by similar methods. For $D'_1 - D_1$,

$$\begin{aligned}
& D'_1 - D_1 \\
=& \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \varepsilon_i \tilde{\tilde{w}}_i(\mathcal{N}_2) \right\}^2 - \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \varepsilon_i \tilde{w}_i(\mathcal{N}_2) \right\}^2 \\
=& \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[w_{ij} - \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right] \right)^2 \\
& - \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[w_{ij} - \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \left\{ \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j} \right\} \right] \right)^2 \\
=& \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right] \right)^2 \\
& - \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \left\{ \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j} \right\} \right] \right)^2 \\
& - \frac{2}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j w_{ij} \right) \\
& \quad \times \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right] \right) \\
& + \frac{2}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j w_{ij} \right) \\
& \quad \times \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j} \left\{ \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j} \right\} \right] \right)
\end{aligned}$$

$=: D_1'^{(1)} - D_1^{(1)} - 2D_1'^{(2)} + 2D_1^{(2)}$.

Now we prove that $D_1'^{(1)} - D_1^{(1)} = o_p\left(\max\left\{1, n_2\delta_n^2\right\}\right)$. Recall that $\hat{e}_j = \varepsilon_j + g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) + \delta_n l(\boldsymbol{X}_j)$. For convenience, define $\hat{\mathcal{G}}^{(ij)} = \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij}\right\}$ and $\mathcal{G}^{(ij)} = \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}$. Then it follows that

$$
\begin{aligned}
& D_1'^{(1)} - D_1^{(1)} \\
=& \frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\hat{\mathcal{G}}^{(ij)}\right)^2 - \frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\mathcal{G}^{(ij)}\right)^2 \\
& + \frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\}\hat{\mathcal{G}}^{(ij)}\right]^2 \\
& - \frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\}\mathcal{G}^{(ij)}\right]^2 \\
& + \frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\hat{\mathcal{G}}^{(ij)}\right\}^2 \\
& - \frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\mathcal{G}^{(ij)}\right\}^2 \\
& + \frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\hat{\mathcal{G}}^{(ij)}\right) \\
& \quad \times \left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\}\hat{\mathcal{G}}^{(ij)}\right] \\
& - \frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\mathcal{G}^{(ij)}\right) \\
& \quad \times \left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\}\mathcal{G}^{(ij)}\right]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\hat{\mathcal{G}}^{(ij)}\right)\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\hat{\mathcal{G}}^{(ij)}\right\}\\
&-\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\mathcal{G}^{(ij)}\right)\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\mathcal{G}^{(ij)}\right\}\\
&+\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\hat{\mathcal{G}}^{(ij)}\right]\\
&\quad\times\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\hat{\mathcal{G}}^{(ij)}\right\}\\
&-\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\mathcal{G}^{(ij)}\right]\\
&\quad\times\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_j)\mathcal{G}^{(ij)}\right\}.
\end{aligned}
$$

We show that the difference between the third and fourth sums satisfies

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\hat{\mathcal{G}}^{(ij)}\right]^2\\
&\quad-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\mathcal{G}^{(ij)}\right]^2\\
=:&\ \hat{\mathcal{L}}-\mathcal{L}=o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)
\end{aligned}
\tag{S2.125}
$$

because we can prove the other terms are $o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$ by similar methods. Rewrite $\mathcal{L}$ as

$$
\begin{aligned}
\mathcal{L}=&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\mathcal{G}^{(ij)}\right]^2\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\right.
\end{aligned}
$$

$$\times \left[\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^{\top} \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}\right] \Bigg)^2$$

$$= \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\} \boldsymbol{\Sigma}^{-1}$$

$$\times \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^{\top}$$

$$\times \boldsymbol{\Sigma}^{-1} \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}. \tag{S2.126}$$

Here $\mathcal{L}$ is the product of $\boldsymbol{\Sigma}^{-1}$, $\frac{1}{n_1}\sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^{\top}$ and $\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}$.

From Condition 4 and (S2.34), we know $\left\|\boldsymbol{\Sigma}^{-1}\right\| = O(1)$ and $\left\|\frac{1}{n_1}\sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^{\top}\right\| = O_p(1)$. Now we derive that

$$\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\} \right\|$$

$$= O_p\left(\max\left\{1, \sqrt{n_2}\delta_n\right\}\right).$$

According to (S2.25),

$$\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0 = \frac{1}{n_2} \sum_{j=n_1+1}^{n} \varepsilon_j \boldsymbol{\Sigma}^{-1} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + \delta_n \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\} + \boldsymbol{\gamma}_{2n},$$

where $\|\boldsymbol{\gamma}_{2n}\| = o_p\left(1/\sqrt{n_2}\right)$. Using Taylor expansion, we have

$$\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}^{\top}$$

$$= \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left(\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_2\right)^{\top} \dot{g}\left(\boldsymbol{\theta}_{(j)}^*, \boldsymbol{X}_j\right) E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}^{\top}$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left(\boldsymbol{\theta}_0-\hat{\boldsymbol{\theta}}_2\right)^{\top}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}^{\top}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left(\boldsymbol{\theta}_0-\hat{\boldsymbol{\theta}}_2\right)^{\top}\left\{\dot{g}\left(\boldsymbol{\theta}^*_{(j)},\boldsymbol{X}_j\right)-\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)\right\}E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}^{\top},
\end{aligned}
\tag{S2.127}
$$

where $\boldsymbol{\theta}^*_{(j)}$ lies between $\hat{\boldsymbol{\theta}}_2$ and $\boldsymbol{\theta}_0$.

The first term in (S2.127) is controlled as

$$
\begin{aligned}
&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left(\boldsymbol{\theta}_0-\hat{\boldsymbol{\theta}}_2\right)^{\top}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}^{\top}\\
=&-\frac{1}{\sqrt{n_2}}\sum_{j_1=n_1+1}^{n}\frac{1}{n_2}\sum_{j_2=n_1+1}^{n}\varepsilon_{j_2}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_2}\right)^{\top}\boldsymbol{\Sigma}^{-1}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_1}\right)E_{\boldsymbol{X}_{j_1}}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j_1}\right\}^{\top}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}^{\top}\boldsymbol{\Sigma}^{-1}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)\\
&\quad\times E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}^{\top}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\boldsymbol{\gamma}_{2n}^{\top}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}^{\top}\\
=&-\frac{1}{\sqrt{n_2}}\sum_{j_2=n_1+1}^{n}\varepsilon_{j_2}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_2}\right)^{\top}\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j_1=n_1+1}^{n}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_{j_1}\right)\\
&\quad\times E_{\boldsymbol{X}_{j_1}}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j_1}\right\}^{\top}\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\delta_n E\left\{\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_1)l(\boldsymbol{X}_1)\right\}^{\top}\boldsymbol{\Sigma}^{-1}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}^{\top}\\
&-\sqrt{n_2}\boldsymbol{\gamma}_{2n}^{\top}\boldsymbol{\Sigma}^{*}_{cov}+o_p\left(1\right)\\
=&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_j\right)^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{*}_{cov}-\sqrt{n_2}\delta_n\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^{*}_{cov}
\end{aligned}
$$

$$
\begin{aligned}
&+ o_p\left(\max\left\{1, \sqrt{n_2}\delta_n\right\}\right) \\
=&O_p\left(\max\left\{1, \sqrt{n_2}\delta_n\right\}\right), \qquad \text{(S2.128)}
\end{aligned}
$$

where $\boldsymbol{\theta}^*_{(j)}$ lies between $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}_2$ and $\boldsymbol{\Sigma}^*_{cov} = E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)^\top w_{12}\right\}$. Next, we explain (S2.128). The reason for the remainder term in (S2.128) is similar to (S2.83), and we omit it. Besides, $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov} = O_p\left(1\right)$ for the following reason: under Condition 9 we know that

$$
E\left[\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)\right\}^4\right] = O\left(1\right),
$$

where $\boldsymbol{\Sigma}^*_{cov} = E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)^\top w_{12}\right\}$. Hence

$$
\begin{aligned}
&E\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^2\right] \\
=&E\left[\varepsilon_1^2\varepsilon_2^2\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)\right\}^2\right] \\
\leq&E\left(\varepsilon_1^4\right)E\left(\varepsilon_2^4\right)E\left[\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)\right\}^4\right] = O\left(1\right).
\end{aligned} \qquad \text{(S2.129)}
$$

Then according to Markov inequality and (S2.129), we obtain that

$$
\begin{aligned}
&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \\
&\times\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1}
\end{aligned}
$$

$$\times \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) + o_p(1) = O_p(1),$$

where $\boldsymbol{\Sigma}_\varepsilon = E\left\{\varepsilon_1^2 \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top\right\}$. Notice that

$$\begin{aligned} O_p(1) =& \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \boldsymbol{\Sigma}^{-1} \\ & \times \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \\ \geq & \left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* \right\|^2 \lambda_{\min}\left(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Sigma}^{-1}\right), \end{aligned} \tag{S2.130}$$

where $\lambda_{\min}\left(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Sigma}^{-1}\right)$ means the smallest eigenvalue of $\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Sigma}^{-1}$.

Under Condition 4, $\lambda_{\min}\left(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Sigma}^{-1}\right)$ is larger than or equal to a positive constant. Then from (S2.130), we derive that

$$\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_j \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_{cov}^* = O_p(1).$$

The proof of (S2.128) is finished.

According to Condition 6,

$$\left\|\dot{g}\left(\boldsymbol{\theta}_{(j)}^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\| \leq \sqrt{p} \left\|\boldsymbol{\theta}_0 - \boldsymbol{\theta}_{(j)}^*\right\| L(\boldsymbol{X}_j) \leq \sqrt{p} \left\|\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_2\right\| L(\boldsymbol{X}_j).$$

Thus, **for the second term** in (S2.127), since $p^3 \delta_n^2 \log n \to 0$ and $\left\|\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_2\right\| = O_p\left(\max\left\{p/n_2, p\delta_n^2\right\}\right)$,

$$\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left(\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_2\right)^\top \left\{\dot{g}\left(\boldsymbol{\theta}_{(j)}^*, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}^\top$$

$$\leq \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\| \boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_2 \right\| \left\| \dot{g}\left(\boldsymbol{\theta}^*_{(j)}, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right) \right\| \left\| E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\} \right\|$$

$$\leq \sqrt{n_2} \left\| \boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_2 \right\|^2 \sqrt{p} L(\boldsymbol{X}_j) C = O_p\left(\max\left\{p^{3/2}/\sqrt{n_2}, p^{3/2}\delta_n^2\sqrt{n_2}\right\}\right) = o_p(1). \tag{S2.131}$$

From (S2.127), (S2.128) and (S2.131), we have

$$\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\} \right\| = O_p(1). \tag{S2.132}$$

Second, We decompose $\hat{\mathcal{L}}$ as the product of three terms, and we prove the differences between these three terms and those in $\mathcal{L}$ are small enough in the sense of L_2 norm.

Similar to (S2.126), we have

$$\begin{aligned}
&\frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} \hat{\mathcal{G}}^{(ij)} \right)^2 \\
=&\frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} \right. \\
&\qquad \left. \times \left[\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right] \right)^2 \\
=&\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} \left[\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right] \hat{\boldsymbol{\Sigma}}^{-1} \\
&\times \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} \\
&\times \left[\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right].
\end{aligned} \tag{S2.133}$$

Now we handle three terms in (S2.133): $\sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top$, $\hat{\boldsymbol{\Sigma}}^{-1}$ and

$$\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} \left[\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right].$$

Deal with $\sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top$. Under Condition 6, we have

$$\left\| \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\| \leq \sqrt{p} \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\| L(\boldsymbol{X}_i).$$

From Lemma 2 we know $\left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\| = O_p\left(\max\left\{ \sqrt{p/n}, \sqrt{p}\delta_n \right\} \right)$. Thus,

$$\begin{aligned}
&\left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top - \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^\top \right\| \\
\leq & 2 \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^\top \right\| \\
& + \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\}^\top \right\| \\
\leq & \frac{2}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 \sqrt{p} \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\| L(\boldsymbol{X}_i) \left\| \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\| + \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i^2 p \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\|^2 L^2(\boldsymbol{X}_i) \\
= & \sqrt{p} O_p\left(\max\left\{ \sqrt{p/n}, \sqrt{p}\delta_n \right\} \right) O_p\left(\sqrt{p}\right) + p O_p\left(\max\left\{ p/n, p\delta_n^2 \right\} \right) = o_p(1),
\end{aligned} \tag{S2.134}$$

where the last equation is because $p^3/n_2 \to 0$ and $p^3 \delta_n^2 \log n \to 0$.

As $p^3 \delta_n^2 \log n \to 0$, we have

$$\begin{aligned}
&\left\| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} \left[\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} \right] \right. \\
& \left. - \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \left\{ g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} E_{\boldsymbol{X}_j} \left\{ \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j} \right\} \right\|
\end{aligned}$$

$$
\begin{aligned}
&\leq \left[\sum_{j=n_1+1}^{n}\left\{g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)-g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)^2\right\}\right]^{1/2} \\
&\quad \times \frac{1}{\sqrt{n_2}}\left[\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij}\right\}-E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}\right\|^2\right]^{1/2} \\
&=O_p\left(\max\left\{\sqrt{p}, \sqrt{pn_2}\delta_n, p\sqrt{pn_2}\delta_n^2\right\}\right) O\left(1/\sqrt{n_2}\right) O_p\left(\max\left\{p, p\sqrt{n_2}\delta_n\right\}\right) \\
&=O_p\left(\max\left\{p^{3/2}/\sqrt{n_2}, p^{3/2}\sqrt{n_2}\delta_n^2, p^{5/2}\sqrt{n_2}\delta_n^3\right\}\right)=o_p(1) \qquad \text{(S2.135)}
\end{aligned}
$$

where

$$
\sum_{j=n_1+1}^{n}\left\{g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)-g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\}^2=O_p\left(\max\left\{p, pn_2\delta_n^2, p^3n_2\delta_n^4\right\}\right) \qquad \text{(S2.136)}
$$

is because of the following. Using the Taylor equation, we have

$$
\begin{aligned}
&\sum_{j=n_1+1}^{n}\left\{g\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)-g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right)\right\}^2 \\
=&\sum_{j=n_1+1}^{n}\left\{\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\dot{g}\left(\tilde{\boldsymbol{\theta}}_{(j)}, \boldsymbol{X}_j\right)\right\}^2 \\
\leq& 2\sum_{j=n_1+1}^{n}\left\{\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\}^2 \\
&+2\sum_{j=n_1+1}^{n}\left[\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_{(j)}, \boldsymbol{X}_j\right)-\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\}\right]^2, \qquad \text{(S2.137)}
\end{aligned}
$$

where $\tilde{\boldsymbol{\theta}}_{(j)}$ lies between $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}_2$. For the first term in (S2.137), we decompose it as

$$
\sum_{j=n_1+1}^{n}\left\{\left(\hat{\boldsymbol{\theta}}_2-\boldsymbol{\theta}_0\right)^{\top}\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\right\}^2
$$

$$= \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)$$

$$= \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top n_1 \boldsymbol{\Sigma} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \tag{S2.138}$$

$$+ \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \left\{ \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)^\top - n_2 \boldsymbol{\Sigma} \right\} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right), \tag{S2.139}$$

where $\boldsymbol{\Sigma} = E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top\right\}$ and $\|\boldsymbol{\Sigma}\| = O(1)$. Recall

$$\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\| = O_p\left(\max\left\{\sqrt{p/n_2}, \sqrt{p}\delta_n\right\}\right).$$

Similar to (S2.35), $\left\|\sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top - n_1\boldsymbol{\Sigma}\right\| = o_p(n_2)$. Then it follows that

$$\begin{aligned}\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top n_2 \boldsymbol{\Sigma} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \leq & n_2 \left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\|^2 \|\boldsymbol{\Sigma}\| \\ = & n_2 O_p\left(\max\left\{p/n_2, p\delta_n^2\right\}\right) O(1) \\ = & O_p\left(\max\left\{p, pn_2\delta_n^2\right\}\right)\end{aligned} \tag{S2.140}$$

and

$$\begin{aligned}&\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \left\{ \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top - n_2 \boldsymbol{\Sigma} \right\} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right) \\ \leq & \left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\|^2 \left\| \sum_{j=n_1+1}^{n} \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top - n_2\boldsymbol{\Sigma} \right\| = o_p\left(\max\left\{p, pn_2\delta_n^2\right\}\right).\end{aligned} \tag{S2.141}$$

According to (S2.138), (S2.140) and (S2.141), we have

$$\sum_{j=n_1+1}^{n} \left\{ \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^\top \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_j) \right\}^2 = O_p\left(\max\left\{p, pn_2\delta_n^2\right\}\right). \tag{S2.142}$$

From Condition 6, we know

$$\left\|\dot{g}\left(\tilde{\boldsymbol{\theta}}_{(j)}, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\| \leq \sqrt{p}\left\|\boldsymbol{\theta}_0 - \tilde{\boldsymbol{\theta}}_{(j)}\right\| L(\boldsymbol{X}_j) \leq \sqrt{p}\left\|\boldsymbol{\theta}_0 - \hat{\boldsymbol{\theta}}_2\right\| L(\boldsymbol{X}_j).$$

Then for the second term in (S2.136),

$$\begin{aligned}
&\sum_{j=n_1+1}^{n}\left[\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right)^{\top}\left\{\dot{g}\left(\tilde{\boldsymbol{\theta}}_{(j)}, \boldsymbol{X}_j\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_j\right)\right\}\right]^2 \\
\leq &\sum_{j=n_1+1}^{n}\left\|\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0\right\|^4 pL^2(\boldsymbol{X}_j) = O_p\left(\max\left\{p^3/n_2, p^3\delta_n^2\right\}\right). \qquad \text{(S2.143)}
\end{aligned}$$

Combining (S2.137), (S2.142) and (S2.143), the equation (S2.136) holds.

As for another term in (S2.135), we have

$$\begin{aligned}
&\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij}\right\} - E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}\right\|^2 \\
\leq &2\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} - \frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) w_{ij}\right\|^2 \\
&+ 2\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) w_{ij} - E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}\right\|^2 \\
= &2\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} - \frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) w_{ij}\right\|^2 \\
&+ 2\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) w_{ij} - E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}\right\|^2
\end{aligned}$$

Here $\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) w_{ij} - E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}\right\|^2 = O_p(p)$. See the discussion about $\boldsymbol{S}_{1n_2}$ in the proof of Lemma 3 for details. Using Cauchy-Schwarz inequality and Condition 6, we have

$$\sum_{j=n_1+1}^{n}\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} - \frac{1}{n_1}\sum_{i=1}^{n_1}\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) w_{ij}\right\|^2$$

$$
\begin{aligned}
&= \sum_{j=n_1+1}^{n} \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\} w_{ij} \right\|^2 \\
&\leq \frac{n_2}{n_1} \sum_{i=1}^{n_1} \left\| \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\|^2 + \sum_{j=n_1+1}^{n} \frac{1}{n_1} \sum_{i=1}^{n_1} w_{ij}^2 \\
&\leq p \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\|^2 \frac{n_2}{n_1} \sum_{i=1}^{n_1} L^2(\boldsymbol{X}_i) + \frac{1}{n_1} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} w_{ij}^2 \\
&= p O_p\left(\max\left\{p/n, p\delta_n^2\right\}\right) O_p\left(n_2\right) + O_p\left(1\right) = O_p\left(\max\left\{p^2, p^2 n_2 \delta_n^2\right\}\right).
\end{aligned}
$$

Then we conclude that

$$
\begin{aligned}
&\sum_{j=n_1+1}^{n} \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\} - E_{\boldsymbol{X}_j}\left\{ \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j} \right\} \right\|^2 \\
&= O_p\left(\max\left\{p^2, p^2 n_2 \delta_n^2\right\}\right).
\end{aligned}
$$

The proof of (S2.135) is finished.

Deal with $\hat{\boldsymbol{\Sigma}}^{-1}$ to prove $\left\| \hat{\boldsymbol{\Sigma}}^{-1} - \boldsymbol{\Sigma}^{-1} \right\| = o_p\left(1\right)$. Similar to the proof of (S2.134), we have

$$
\begin{aligned}
&\left\| \frac{1}{n} \sum_{i=1}^{n} \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top - \frac{1}{n} \sum_{i=1}^{n} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^\top \right\| \\
\leq & 2 \left\| \frac{1}{n} \sum_{i=1}^{n} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\} \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right)^\top \right\| \\
& + \left\| \frac{1}{n} \sum_{i=1}^{n} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) - \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\}^\top \right\| \\
\leq & \frac{2}{n} \sum_{i=1}^{n} \sqrt{p} \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\| L(\boldsymbol{X}_i) \left\| \dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_i\right) \right\| + \frac{1}{n} \sum_{i=1}^{n} p \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\|^2 L^2(\boldsymbol{X}_i) \\
= & \sqrt{p} O_p\left(\max\left\{ \sqrt{p/n}, \sqrt{p}\delta_n \right\} \right) O_p\left(\sqrt{p}\right) + p O_p\left(\max\left\{p/n, p\delta_n^2\right\}\right)
\end{aligned}
$$

$$=o_p\left(1\right),$$

and similar to (S2.33), we have $\left\|\frac{1}{n}\sum_{i=1}^{n}\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_i\right)\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_i\right)^{\top}-\boldsymbol{\Sigma}\right\| = o_p(1)$. So $\left\|\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma}\right\| = o_p\left(1\right)$. For $\hat{\boldsymbol{\Sigma}}^{-1}-\boldsymbol{\Sigma}^{-1}$,

$$\begin{aligned}\left\|\hat{\boldsymbol{\Sigma}}^{-1}-\boldsymbol{\Sigma}^{-1}\right\| =& \left\|\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}-\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}\right\| \\ =& \left\|\hat{\boldsymbol{\Sigma}}^{-1}\left(\boldsymbol{\Sigma}-\hat{\boldsymbol{\Sigma}}\right)\boldsymbol{\Sigma}^{-1}\right\| \\ \leq& \left\|\hat{\boldsymbol{\Sigma}}^{-1}\right\|\left\|\boldsymbol{\Sigma}-\hat{\boldsymbol{\Sigma}}\right\|\left\|\boldsymbol{\Sigma}^{-1}\right\| \\ =& \lambda_{\min}^{-1}\left(\hat{\boldsymbol{\Sigma}}\right)\left\|\boldsymbol{\Sigma}-\hat{\boldsymbol{\Sigma}}\right\|\lambda_{\min}^{-1}\left(\boldsymbol{\Sigma}\right). \end{aligned} \tag{S2.144}$$

From Condition 4, we know that $\lambda_{\min}^{-1}\left(\boldsymbol{\Sigma}\right)\leq a^{-1}$, $\lambda_{\min}\left(\boldsymbol{\Sigma}\right)\geq a$. If we can prove that $\lambda_{\min}\left(\hat{\boldsymbol{\Sigma}}\right)-\lambda_{\min}\left(\boldsymbol{\Sigma}\right)=o_p\left(1\right)$, and then $\lambda_{\min}\left(\hat{\boldsymbol{\Sigma}}\right)\geq\frac{a}{2}$ holds with the probability converging to 1. Thus, we can infer that

$$\left\|\hat{\boldsymbol{\Sigma}}^{-1}-\boldsymbol{\Sigma}^{-1}\right\|\leq\frac{4}{a^2}\left\|\boldsymbol{\Sigma}-\hat{\boldsymbol{\Sigma}}\right\|=o_p\left(1\right).$$

According to equation (5.1) in Zhu and Fang (1996), denoting $\boldsymbol{b}_{\min}(\boldsymbol{A})$ as the unit eigenvector corresponding to the smallest eigenvalue of $\boldsymbol{A}$, we have

$$\left\{\lambda_{\min}\left(\hat{\boldsymbol{\Sigma}}\right)-\lambda_{\min}\left(\boldsymbol{\Sigma}\right)\right\}=\boldsymbol{b}_{\min}(\boldsymbol{\Sigma})^{\top}\left(\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma}\right)\boldsymbol{b}_{\min}(\boldsymbol{\Sigma})+R_{\min}\left(\boldsymbol{\Sigma}^{*}\right),$$

where

$$R_{\min}\left(\boldsymbol{\Sigma}^{*}\right)=\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^{*})^{\top}\left(\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma}\right)\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^{*})-\boldsymbol{b}_{\min}(\boldsymbol{\Sigma})^{\top}\left(\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma}\right)\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}),$$

and $\boldsymbol{\Sigma}^{*}$ lies between $\boldsymbol{\Sigma}$ and $\hat{\boldsymbol{\Sigma}}$. Since $\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^{*})$ is the unit eigenvector, we

have $\|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)\| = 1$.

$$
\begin{aligned}
|R_{\min}(\boldsymbol{\Sigma}^*)| = & \left|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)^\top \left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right) \boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*) - \boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)^\top \left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right) \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})\right. \\
& \left. + \boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)^\top \left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right) \boldsymbol{b}_{\min}(\boldsymbol{\Sigma}) - \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})^\top \left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right) \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})\right| \\
\leq & \left|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)^\top \left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right) \{\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*) - \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})\}\right| \\
& + \left|\left\{\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)^\top - \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})^\top\right\} \left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right) \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})\right| \\
\leq & \left\|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)^\top\right\| \left\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right\| \left\|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*) - \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})\right\| \\
& + \left\|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma}^*)^\top - \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})^\top\right\| \left\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right\| \left\|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma})\right\| \\
\leq & 2\left\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right\| + 2\left\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right\| \\
= & 4\left\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right\|.
\end{aligned}
$$

Then we conclude that

$$
\begin{aligned}
& \left\{\lambda_{\min}\left(\hat{\boldsymbol{\Sigma}}\right) - \lambda_{\min}(\boldsymbol{\Sigma})\right\} \\
= & \boldsymbol{b}_{\min}(\boldsymbol{\Sigma})^\top \left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right) \boldsymbol{b}_{\min}(\boldsymbol{\Sigma}) + R_{\min}(\boldsymbol{\Sigma}^*) \\
\leq & \|\boldsymbol{b}_{\min}(\boldsymbol{\Sigma})\|^2 \left\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right\| + 4\left\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right\| = o_p(1).
\end{aligned}
$$

Hence, from (S2.144) we have

$$
\left\|\hat{\boldsymbol{\Sigma}}^{-1} - \boldsymbol{\Sigma}^{-1}\right\| = o_p(1). \tag{S2.145}
$$

Right now, we have proved the three terms in $\hat{\mathcal{L}}$ are close enough to the three terms in $\mathcal{L}$ in the sense of norm. Then from (S2.126), (S2.132), (S2.133), (S2.134), (S2.135), (S2.145), $\left\|\boldsymbol{\Sigma}^{-1}\right\| = O(1)$ and

$\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_i\right)\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_i\right)^\top\right\| = O_p\left(1\right)$, using Lemma 15, we have

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\hat{\mathcal{G}}^{(ij)}\right]^2\\
&-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\left[\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\mathcal{G}^{(ij)}\right]^2\\
=&o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right), \qquad \text{(S2.146)}
\end{aligned}$$

as the left-hand side divided by $n_2\delta_n^2$ satisfies the conditions of Lemma 15 when $\lim_{n\to\infty} n_2\delta_n^2 \to \infty$, otherwise, the left-hand size naturally satisfies the conditions when $n_2\delta_n^2 = O\left(1\right)$.

Similarly, we can prove that $D'^{(1)}_1 - D^{(1)}_1 = o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$ and $D'^{(2)}_1 - D^{(2)}_1 = o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$. Hence $D'_1 - D_1 = o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$. We omit the discussions of $D'_2 - D_2 = o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$ to $D'_6 - D_6 = o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$ as they are similar. Thus, we arrive at the conclusion of *Step 3*: $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2 - \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$. *Step 3* is finished. *Step 4*. Prove $\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2 - \left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2 = o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$. Obviously,

$$\begin{aligned}
&\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2 - \left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2\\
=&2\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\\
&+\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}
\end{aligned}$$

$$\times \left\{ \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{w}_i(\mathcal{N}_2) \right\}. \tag{S2.147}$$

From *Step 2* we know that $\frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{w}_i(\mathcal{N}_2) = O_p\left(\max\left\{1, \sqrt{n_2}\delta_n\right\}\right)$. So according to (S2.147) we just need to show

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{w}_i(\mathcal{N}_2) = o_p\left(\max\left\{1, \sqrt{n_2}\delta_n\right\}\right).$$

For this term, recall that $\hat{\mathcal{G}}^{(ij)} = \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i\right) w_{ij} \right\}$ and

$\mathcal{G}^{(ij)} = \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0, \boldsymbol{X}_1\right) w_{1j}\right\}$. Thus $\tilde{\tilde{w}}_i(\mathcal{N}_2) - \tilde{w}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left(\mathcal{G}^{(ij)} - \hat{\mathcal{G}}^{(ij)}\right)$. Note that $\hat{e}_j = \varepsilon_j + g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) + \delta_n l(\boldsymbol{X}_j)$. Then we have

$$\begin{aligned}
&\frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{w}_i(\mathcal{N}_2) \\
=&\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \varepsilon_i + g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) + \delta_n l(\boldsymbol{X}_i) \right\} \\
&+ \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left(\mathcal{G}^{(ij)} - \hat{\mathcal{G}}^{(ij)}\right) \\
=&\frac{1}{n_1\sqrt{n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j \left(\mathcal{G}^{(ij)} - \hat{\mathcal{G}}^{(ij)}\right) \\
&+ \frac{1}{n_1\sqrt{n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \delta_n^2 l(\boldsymbol{X}_i) l(\boldsymbol{X}_j) \left(\mathcal{G}^{(ij)} - \hat{\mathcal{G}}^{(ij)}\right) \\
&+ \frac{1}{n_1\sqrt{n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right) \right\} \left\{ g(\boldsymbol{\theta}_0, \boldsymbol{X}_j) - g\left(\hat{\boldsymbol{\theta}}_2, \boldsymbol{X}_j\right) \right\} \\
&\quad \times \left(\mathcal{G}^{(ij)} - \hat{\mathcal{G}}^{(ij)}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\left(\mathcal{G}^{(ij)}-\hat{\mathcal{G}}^{(ij)}\right)\\
&+\frac{1}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\varepsilon_j\left(\mathcal{G}^{(ij)}-\hat{\mathcal{G}}^{(ij)}\right)\\
&+\frac{1}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\delta_n\varepsilon_i l(\boldsymbol{X}_j)\left(\mathcal{G}^{(ij)}-\hat{\mathcal{G}}^{(ij)}\right)\\
&+\frac{1}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_i)\varepsilon_j\left(\mathcal{G}^{(ij)}-\hat{\mathcal{G}}^{(ij)}\right)\\
&+\frac{1}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\delta_n l(\boldsymbol{X}_i)\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\left(\mathcal{G}^{(ij)}-\hat{\mathcal{G}}^{(ij)}\right)\\
&+\frac{1}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\delta_n l(\boldsymbol{X}_j)\left(\mathcal{G}^{(ij)}-\hat{\mathcal{G}}^{(ij)}\right).
\end{aligned}
\tag{S2.148}
$$

We here only give a brief discussion on the third sum to derive

$$
\begin{aligned}
&\frac{1}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\left\{g(\boldsymbol{\theta}_0,\boldsymbol{X}_j)-g\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}\\
&\quad\times\left(\mathcal{G}^{(ij)}-\hat{\mathcal{G}}^{(ij)}\right)\\
=&o_p\left(\max\left\{1,\sqrt{n_2}\delta_n\right\}\right),
\end{aligned}
\tag{S2.149}
$$

as the other sums can be similarly handled. Further, by checking the proof for (S2.125), we can easily see the similarity when we deal with this sum. Therefore, we omit the details. *Step 4* is finished.

From *Step 1* to *Step 4* we conclude that

$$
\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2
$$

$$
\begin{aligned}
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2-\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2-\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i(\mathcal{N}_2)\right\}^2+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)\\
=&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]+o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right).
\end{aligned}
$$

Eventually, under the null hypothesis with $\delta_n = 0$, according to (S2.72) in the proof of Lemma 9, we have that

$$
\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}-1=o_p(1).
$$

Under the local alternative hypothesis with $p^3\delta_n^2\to 0$, according to Lemma 14, we derive that

$$
\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+\sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}-1=o_p(1).
$$

□

Corollary 2. *Suppose that Conditions 1-11 hold. Under the null hypothesis and the local alternative hypothesis in (S2.7) with $p^3\delta_n^2\to 0$, recalling $e_i = \varepsilon_i + \delta_n l(\boldsymbol{X}_i)$, we have*

$$
\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{e_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}e_i\tilde{w}_i(\mathcal{N}_2)\right\}^2}{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}\xrightarrow{p}1,
$$

which is mentioned in the main context.

Proof of Corollary 2. The proof is similar to that for Lemma 16. Thus we omit the details here. □

Lemma 17. *Suppose that Conditions 1-10 hold. Under the global alternative hypothesis in (S2.7) with a fixed nonzero δ_n, recall $m(\boldsymbol{X}) = E(Y|\boldsymbol{X})$ and $l(\boldsymbol{X}) = m(\boldsymbol{X}) - g(\boldsymbol{\theta}^*, \boldsymbol{X})$. If $p^3 \log n_1/n_1 \to 0$ and $p^3 \log n_2/n_2 \to 0$, then as $n_1 \to \infty$ and $n_2 \to \infty$,*

$$
\begin{aligned}
&\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \left\{ \hat{e}_i \tilde{w}_i^{(0)}(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{w}_i^{(0)}(\mathcal{N}_2) \right\}^2 \\
=& E\left[\left\{ \varepsilon_1^2 + l^2(\boldsymbol{X}_1) \right\} E_{\boldsymbol{X}_1}^2 \left\{ l(\boldsymbol{X}_2) w_{12} \right\} \right] - E^2 \left\{ l(\boldsymbol{X}_1) l(\boldsymbol{X}_2) w_{12} \right\} \\
&+ E\left\{ l(\boldsymbol{X}_2) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{12} \right\}^\top \boldsymbol{\Sigma}_*^{-1} \left(\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_l \right) \boldsymbol{\Sigma}_*^{-1} E\left\{ l(\boldsymbol{X}_2) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{12} \right\} \\
&- 2E\left[\left\{ \varepsilon_1^2 + l^2(\boldsymbol{X}_1) \right\} l(\boldsymbol{X}_2) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{12} \right]^\top \boldsymbol{\Sigma}_*^{-1} E\left\{ l(\boldsymbol{X}_2) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{12} \right\} \\
&+ o_p(1),
\end{aligned}
$$

where $\boldsymbol{\Sigma}_* = E\left\{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top \right\}$, $\boldsymbol{\Sigma}_\varepsilon = E\left\{ \varepsilon_1^2 \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top \right\}$, $\boldsymbol{\Sigma}_l = E\left\{ l^2(\boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top \right\}$, *and*
$\tilde{w}_i^{(0)}(\mathcal{N}_2) = \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j \left[w_{ij} - \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}_*^{-1} E_{\boldsymbol{X}_j} \{ \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{1j} \} \right]$.

Proof of Lemma 17. Firstly, analyse $\tilde{w}_1^{(0)}(\mathcal{N}_2)$. Before our discussion, we analyse two important components of $\tilde{w}_1^{(0)}(\mathcal{N}_2)$. Under the global alternative hypothesis, using case (b) in Lemma 3, we have

$$
\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \hat{e}_j w_{1j}
$$

$$
\begin{aligned}
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j w_{1j}+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)-\dot{g}\left(\hat{\boldsymbol{\theta}}_2,\boldsymbol{X}_j\right)\right\}w_{1j}\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)w_{1j}\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\right\}\right]\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\right\}\right]-\mathcal{S}_1,
\end{aligned}
\tag{S2.150}
$$

where the remainder term $\mathcal{S}_1$ safisfies $E_{\mathcal{N}_2}(\mathcal{S}_1^4)=o_p(1)$. By a similar discussion on the remainder term in equation (S2.52), we can derive that under Condition 4,

$$
\begin{aligned}
&\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j}\}\\
=&\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}E_{\mathcal{N}_2}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j w_{1j}\right\}\\
=&\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}E_{\mathcal{N}_2}\Bigg\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\right.\right.\\
&\left.\left.\times E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\right\}\right]\right)\Bigg\}\\
&+\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}E_{\mathcal{N}_2}\Bigg\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\right.\\
&\left.\times E_{\boldsymbol{X}_1}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\right\}\right]\Bigg\}-\tilde{\mathcal{S}}_1,
\end{aligned}
$$

where $\tilde{\mathcal{S}}_1=\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}E_{\mathcal{N}_2}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\mathcal{S}_1\}$. Thus, we can decompose

$\tilde{w}_1^{(0)}(\mathcal{N}_2)$ as

$$
\begin{aligned}
&\tilde{w}_1^{(0)}(\mathcal{N}_2)\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j}\}\right]\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\}\right]\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\}\right]-\mathcal{S}_1+\tilde{\mathcal{S}}_1\\
&-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}E_{\mathcal{N}_2}\Bigg\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\Bigg(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\big[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\\
&\quad\times E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\}\big]\Bigg)\Bigg\}\\
&-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}E_{\mathcal{N}_2}\Bigg\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)\big[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\\
&\quad\times E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\}\big]\Bigg\}\\
=&\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\}\right]\\
&+\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)\left[w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\}\right]-\mathcal{S}_1+\tilde{\mathcal{S}}_1\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}\\
&\quad\times E_{\boldsymbol{X}_j}\left[\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\left\{w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\right\}\right]\\
&-\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\boldsymbol{\Sigma}_*^{-1}\\
&\quad\times E_{\boldsymbol{X}_j}\left[\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\left\{w_{1j}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\right\}\right]
\end{aligned}
$$

$$=: \sqrt{n_2}S_{11} + \sqrt{n_2}S_{12} - \mathcal{S}_1 + \tilde{\mathcal{S}}_1 + \sqrt{n_2}S_{13} + \sqrt{n_2}S_{14}^*.$$

By replacing 1 with i, we can define S_{i1}, S_{i2}, S_{i3}, S_{i4}^*, $\mathcal{S}_i$ and $\tilde{\mathcal{S}}_i$. For S_{i4}^*, we have

$$\begin{aligned}
S_{i4}^* =& -\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}_*^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{1j}\right\} \\
&+ \frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}_*^{-1} \\
&\quad \times E_{\boldsymbol{X}_j}\left[\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j)^\top \boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)w_{12}\right\}\right] \\
=& -\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}_*^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{1j}\right\} \\
&+ \frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}_*^{-1}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_j) \\
=:& S_{i4} + S_{i5},
\end{aligned}$$

where $\boldsymbol{\Sigma}_\varepsilon = E\left\{\varepsilon_1^2\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top\right\}$.

In the following, we decompose the proof into two steps.

Step 1. Prove

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2 \\
=&E\left[\left\{\varepsilon_1^2 + l^2(\boldsymbol{X}_1)\right\}E_{\boldsymbol{X}_1}^2\left\{l(\boldsymbol{X}_2)w_{12}\right\}\right] \\
&+ E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\}^\top \boldsymbol{\Sigma}_*^{-1}\left(\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_l\right)\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\} \\
&- 2E\left[\left\{\varepsilon_1^2 + l^2(\boldsymbol{X}_1)\right\}l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right]^\top \boldsymbol{\Sigma}_*^{-1}
\end{aligned}$$

$$\times E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\} + o_p(1).$$

Decompose $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2$ as

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
=&\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left[\left\{\varepsilon_i + l(\boldsymbol{X}_i) + g\left(\boldsymbol{\theta}^*, \boldsymbol{X}_i\right) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right]^2 \\
=&\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 + \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
&+ \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left[\left\{g(\boldsymbol{\theta}^*, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right]^2 \\
&+ \frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\varepsilon_i l(\boldsymbol{X}_i)\left\{\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
&+ \frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\varepsilon_i\left\{g(\boldsymbol{\theta}^*, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\left\{\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
&+ \frac{2}{n_1 n_2}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)\left\{g(\boldsymbol{\theta}^*, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\left\{\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
=:& G_1 + G_2 + G_3 + G_4 + G_5 + G_6. \qquad \text{(S2.151)}
\end{aligned}$$

Now we start to analyse G_i, $i = 1, 2, 3$, and G_i, $i = 4, 5, 6$ can be handled by Cauchy-Schwarz inequality. G_1 can be further decomposed as

$$G_1 = \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\left(S_{i1} + S_{i2} + S_{i3} + S_{i4} + S_{i5} - \frac{1}{\sqrt{n_2}}\mathcal{S}_i + \frac{1}{\sqrt{n_2}}\tilde{\mathcal{S}}_i\right)\right\}^2.$$

Here, we can decompose G_1 as the sum of many terms. However, only the terms $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i2}^2$, $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i4}^2$ and $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i2}S_{i4}$ do not converge to zero in probability. We will show that in the following. For $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i2}^2$,

we have

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i2}^2 \\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\frac{1}{n_2^2}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_1})w_{ij_1}l(\boldsymbol{X}_{j_2})w_{ij_2} \\
&+\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\frac{1}{n_2^2}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_1})\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_{j_1})^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{i2}\right\} \\
&\quad\times l(\boldsymbol{X}_{j_2})\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_{j_2})^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{i2}\right\} \\
&-\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\frac{1}{n_2^2}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_1})w_{ij_1}l(\boldsymbol{X}_{j_2})\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_{j_2})^\top\boldsymbol{\Sigma}^{-1} \\
&\quad\times E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{i2}\right\} \\
=&E\left[\varepsilon_1^2E_{\boldsymbol{X}_1}^2\left\{l(\boldsymbol{X}_2)w_{12}\right\}\right]+o_p(1)=O_p(1),
\end{aligned}
\tag{S2.152}
$$

where the terms from line 3 to line 5 are $o_p(1)$, since under the global alternative hypothesis, $E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\right\}=\mathbf{0}$, which has been proved in (S2.15). Besides, the last equation in (S2.152) is because $E\left[\varepsilon_1^2E_{\boldsymbol{X}_1}^2\left\{l(\boldsymbol{X}_2)w_{12}\right\}\right]\le E^{1/2}\left(\varepsilon_1^4\right)E^{1/2}\left\{l^4(\boldsymbol{X}_1)\right\}E^{1/2}\left(w_{12}^4\right)=O(1)$.

For $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2S_{i4}^2$, we have

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i4}^2 \\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\frac{1}{n_2^2}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_1})l(\boldsymbol{X}_{j_2})E_{\boldsymbol{X}_{j_1}}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_1}\right\}^\top\boldsymbol{\Sigma}_*^{-1} \\
&\times\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}_*^{-1}E_{\boldsymbol{X}_{j_2}}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{1}{n_2}\sum_{j_1=n_1+1}^{n} l(\boldsymbol{X}_{j_1})E_{\boldsymbol{X}_{j_1}}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_1}\}^\top\boldsymbol{\Sigma}_*^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top \\
&\times\boldsymbol{\Sigma}_*^{-1}\frac{1}{n_2}\sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_2})E_{\boldsymbol{X}_{j_2}}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_2}\} \\
=&E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\}^\top\boldsymbol{\Sigma}_*^{-1}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\}+o_p(1) \\
=&O_p(1). \qquad \text{(S2.153)}
\end{aligned}
$$

where $\boldsymbol{\Sigma}_\varepsilon = E\left\{\varepsilon_1^2\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\right\}$. Note that in the last line of (S2.153), $\|E\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\}\| = O(1)$ for the following reason: under Conditions 7 and 10, using Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
&\|E\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\}\|^2 \\
=&\sup_{\|\boldsymbol{a}\|=1}\left[\boldsymbol{a}^\top E\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\}\right]^2 \\
\le&\sup_{\|\boldsymbol{a}\|=1} E\left[\left\{\boldsymbol{a}^\top l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\right\}^2\right]E\left(w_{12}^2\right) \\
\le&\sup_{\|\boldsymbol{a}\|=1}\boldsymbol{a}^\top E\left\{l^2(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\right\}\boldsymbol{a}\sqrt{C_1} \\
=&\left\|E\left\{l^2(\boldsymbol{X}_2)\right\}E\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)^\top\right\}\right\| \\
=&E\left\{l^2(\boldsymbol{X}_2)\right\}\|\boldsymbol{\Sigma}_*\| = O(1). \qquad \text{(S2.154)}
\end{aligned}
$$

For $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i5}^2$, recall $E\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\} = \mathbf{0}$ in (S2.15). Then we have

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i5}^2 \\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\frac{1}{n_2^2}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_1})l(\boldsymbol{X}_{j_2})\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_{j_1})^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{cov}^*\boldsymbol{\Sigma}_*^{-1}
\end{aligned}
$$

$$
\begin{aligned}
&\times \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \boldsymbol{\Sigma}_*^{-1}\boldsymbol{\Sigma}_{cov}^*\boldsymbol{\Sigma}^{-1}\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_2}) \\
=&\frac{1}{n_2}\sum_{j_1=n_1+1}^{n} l(\boldsymbol{X}_{j_1})\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_1})^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{cov}^*\boldsymbol{\Sigma}_*^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \\
&\times \boldsymbol{\Sigma}_*^{-1}\boldsymbol{\Sigma}_{cov}^*\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j_2=n_1+1}^{n} l(\boldsymbol{X}_{j_2})\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_{j_2}) \\
=&E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}_*^{-1}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\right\} + o_p(1) \\
=&o_p(1),
\end{aligned}
$$

where $\boldsymbol{\Sigma}_{cov}^* = E\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_2)^\top w_{12}\right\}$.

Through similar or easier discussions, we can prove that $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i1}^2 = o_p(1)$, $\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i3}^2 = o_p(1)$,

$$
\begin{aligned}
\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i2}S_{i4} =& - E\left\{\varepsilon_1^2 l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\}^\top \boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\} \\
&+ o_p(1) = O_p(1),
\end{aligned} \tag{S2.155}
$$

and the other terms in (S2.150) converge to zero. Hence, from (S2.150), (S2.152), (S2.153) and (S2.155), we infer that

$$
\begin{aligned}
G_1 =&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i2}^2 + \frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i4}^2 + \frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i^2 S_{i2}S_{i4} + o_p(1) \\
=&E\left[\varepsilon_1^2 E_{\boldsymbol{X}_1}^2\left\{l(\boldsymbol{X}_2)w_{12}\right\}\right] \\
&+ E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\}^\top \boldsymbol{\Sigma}_*^{-1}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\}
\end{aligned}
$$

$$- 2E\left\{\varepsilon_1^2 l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\}^\top \boldsymbol{\Sigma}_*^{-1} E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\} + o_p(1)$$

$$= O_p(1). \tag{S2.156}$$

By replacing "ε_i" in $G_1 = \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i \frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2$ with "$l(\boldsymbol{X}_i)$", we can prove that

$$\begin{aligned} G_2 =& \frac{1}{n_1}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\ =& E\left[l^2(\boldsymbol{X}_1)E_{\boldsymbol{X}_1}^2\left\{l(\boldsymbol{X}_2)w_{12}\right\}\right] \\ &+ E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\}^\top \boldsymbol{\Sigma}_*^{-1}\boldsymbol{\Sigma}_l\boldsymbol{\Sigma}_*^{-1} E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\} \\ &- 2E\left\{l^2(\boldsymbol{X}_1)l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\}^\top \boldsymbol{\Sigma}_*^{-1} E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)w_{12}\right\} + o_p(1) \\ =& O_p(1), \end{aligned} \tag{S2.157}$$

where $\boldsymbol{\Sigma}_l = E\left\{l^2(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)^\top\right\}$.

For G_3,

$$\begin{aligned} G_3 =& \frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{g(\boldsymbol{\theta}^*, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}\right. \\ &\left.\times\left(S_{i1} + S_{i2} + S_{i3} + S_{i4} + S_{i5} - \frac{1}{\sqrt{n_2}}\mathcal{S}_i + \frac{1}{\sqrt{n_2}}\tilde{\mathcal{S}}_i\right)\right]^2. \end{aligned}$$

As an example, we just prove $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}^*, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}^2 S_{i4}^2 = o_p(1)$, as we can prove that the other terms in G_3 converge to zero by similar or easier methods.

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}^*, \boldsymbol{X}_i) - g\left(\hat{\boldsymbol{\theta}}_1, \boldsymbol{X}_i\right)\right\}^2 S_{i4}^2$$

$$
\begin{aligned}
=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}^*,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}^2\frac{1}{n_2^2}\sum_{j_1=n_1+1}^{n}\sum_{j_2=n_1+1}^{n}l(\boldsymbol{X}_{j_1})l(\boldsymbol{X}_{j_2})\\
&\times E_{\boldsymbol{X}_{j_1}}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_1}\}^\top\boldsymbol{\Sigma}_*^{-1}\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}_*^{-1}E_{\boldsymbol{X}_{j_2}}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_2}\}\\
=&\frac{1}{n_2}\sum_{j_1=n_1+1}^{n}l(\boldsymbol{X}_{j_1})E_{\boldsymbol{X}_{j_1}}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_1}\}^\top\boldsymbol{\Sigma}_*^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}^*,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}^2\\
&\times\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}_*^{-1}\frac{1}{n_2}\sum_{j_2=n_1+1}^{n}l(\boldsymbol{X}_{j_2})E_{\boldsymbol{X}_{j_2}}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j_2}\}.
\end{aligned}
\tag{S2.158}
$$

From (S2.154), we know that

$$
\begin{aligned}
&\left\|\frac{1}{n_2}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)E_{\boldsymbol{X}_j}\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{1j}\}\right\|\\
=&\left\|E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\}\right\|+o_p\left(1\right)=O_p\left(1\right).
\end{aligned}
$$

Under Condition 4, $\left\|\boldsymbol{\Sigma}_*^{-1}\right\|=O\left(1\right)$. Note that

$$
\begin{aligned}
&\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}^*,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}^2\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top\right\|\\
=&\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\left(\boldsymbol{\theta}^*-\hat{\boldsymbol{\theta}}_1\right)^\top\dot{g}(\boldsymbol{\theta}^*_{(i)},\boldsymbol{X}_i)\right\}^2\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top\right\|\\
\leq&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\|\boldsymbol{\theta}^*-\hat{\boldsymbol{\theta}}_1\right\|^2\left\|\dot{g}(\boldsymbol{\theta}^*_{(i)},\boldsymbol{X}_i)\right\|\left\|\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\right\|^2\\
\leq&O_p\left(p/n_1\right)O_p\left(\sqrt{p}\right)O_p\left(p\right)=o_p\left(1\right),
\end{aligned}
$$

where $\boldsymbol{\theta}^*_{(i)}$ lies between $\boldsymbol{\theta}^*$ and $\hat{\boldsymbol{\theta}}_1$. Hence, from (S2.158) we know $\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{g(\boldsymbol{\theta}^*,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}^2S_{i4}^2=o_p\left(1\right)$. Similarly, the other terms in G_3 are $o_p\left(1\right)$. We can prove G_4, G_5 and G_6 are $o_p\left(1\right)$ by Cauchy-Schwarz

inequality. Thus, from (S2.151), (S2.156) and (S2.157) we infer that

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2 \\
=&E\left[\left\{\varepsilon_1^2+l^2(\boldsymbol{X}_1)\right\}E_{\boldsymbol{X}_1}^2\left\{l(\boldsymbol{X}_2)w_{12}\right\}\right] \\
&+E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\}^\top\boldsymbol{\Sigma}_*^{-1}\left(\boldsymbol{\Sigma}_\varepsilon+\boldsymbol{\Sigma}_l\right)\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\} \\
&-2E\left[\left\{\varepsilon_1^2+l^2(\boldsymbol{X}_1)\right\}l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right]^\top\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\} \\
&+o_p\left(1\right).
\end{aligned}
$$

Step 1 is finished.

Step 2. Prove

$$
\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2=E^2\left\{l\left(\boldsymbol{X}_1\right)l\left(\boldsymbol{X}_2\right)w_{12}\right\}+o_p\left(1\right).
$$

Decompose this term as

$$
\begin{aligned}
&\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2 \\
=&\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i+l(\boldsymbol{X}_i)+g\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right]^2 \\
=&\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2+\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
&+\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\left\{g(\boldsymbol{\theta}^*,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right]^2 \\
&+\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i l(\boldsymbol{X}_i)\left\{\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 \\
&+\frac{2}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\left\{g(\boldsymbol{\theta}^*,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\left\{\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2
\end{aligned}
$$

$$+\frac{2}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)\left\{g(\boldsymbol{\theta}^*,\boldsymbol{X}_i)-g\left(\hat{\boldsymbol{\theta}}_1,\boldsymbol{X}_i\right)\right\}\left\{\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2. \quad \text{(S2.159)}$$

We show that for the second term in (S2.159),

$$\left\{\frac{1}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2 = E^2\left\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}+o_p(1) \quad \text{(S2.160)}$$

as we can show the other terms are $o_p(1)$ in (S2.159) by similar methods.

For $\left\{\frac{1}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2$, we have

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}(\mathcal{N}_2)\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)\left(S_{i1}+S_{i2}+S_{i3}+S_{i4}+S_{i5}-\frac{1}{\sqrt{n_2}}\mathcal{S}_i+\frac{1}{\sqrt{n_2}}\tilde{\mathcal{S}}_i\right).
\end{aligned} \quad \text{(S2.161)}$$

Note that

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)S_{i2}\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\left[w_{ij}-\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)^\top\boldsymbol{\Sigma}^{-1}E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{i2}\right\}\right]\\
=&\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_i)l(\boldsymbol{X}_j)w_{ij}\\
&-\frac{1}{n_1}\sum_{i=1}^{n_1} l(\boldsymbol{X}_i)E_{\boldsymbol{X}_i}\left\{\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{i2}\right\}^\top\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_j)\\
=&E\left\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}-E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_2)w_{12}\right\}\boldsymbol{\Sigma}^{-1}E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\right\}\\
&+o_p(1)\\
=&E\left\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}+o_p(1),
\end{aligned}$$

where the last equation is because $E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\right\} = 0$.

Similarly,

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} l\left(\boldsymbol{X}_i\right) S_{i4} \\
=&\frac{1}{n_1}\sum_{i=1}^{n_1} l\left(\boldsymbol{X}_i\right) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_i)^\top \frac{1}{n_2}\sum_{j=n_1+1}^{n} l\left(\boldsymbol{X}_j\right) \boldsymbol{\Sigma}_*^{-1} E_{\boldsymbol{X}_j}\left\{\dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{1j}\right\} \\
=&E\left\{l\left(\boldsymbol{X}_1\right) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}_*^{-1} E\left\{l\left(\boldsymbol{X}_2\right) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1) w_{12}\right\} + o_p\left(1\right) = o_p\left(1\right),
\end{aligned}$$

where the last equation is derived from the fact that $E\left\{l\left(\boldsymbol{X}_1\right) \dot{g}(\boldsymbol{\theta}^*, \boldsymbol{X}_1)\right\} = 0$. And we can prove the other terms in (S2.161) are $o_p\left(1\right)$ by easier methods. Hence, (S2.160) holds. Similarly, the other terms in (S2.159) are $o_p\left(1\right)$. Then it follows that

$$\left\{\frac{1}{n_1}\sum_{i=1}^{n_1} \hat{e}_i \frac{1}{\sqrt{n_2}} \tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2 = E^2\left\{l\left(\boldsymbol{X}_1\right) l(\boldsymbol{X}_2) w_{12}\right\} + o_p\left(1\right).$$

The proof of *Step 2* is finished.

Combining the results of *Step 1* and *Step 2*, we know the conclusion of Lemma 17 holds. □

Lemma 18. *Suppose that Conditions 1-10 hold. Under the global alternative hypothesis in (S2.7) with a fixed nonzero δ_n, recall $m(\boldsymbol{X}) = E(Y|\boldsymbol{X})$ and $l(\boldsymbol{X}) = m(\boldsymbol{X}) - g(\boldsymbol{\theta}^*, \boldsymbol{X})$. If $p^3 \log n_1/n_1 \to 0$ and $p^3 \log n_2/n_2 \to 0$, then when $n_1 \to \infty$ and $n_2 \to \infty$, we have that*

$$\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left\{\hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2$$

$$=\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{w}_i^{(0)}(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{w}_i^{(0)}(\mathcal{N}_2)\right\}^2+o_p(1),$$

where

$$\tilde{w}_i^{(0)}(\mathcal{N}_2)=\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j\left[w_{ij}-\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)^\top\boldsymbol{\Sigma}_*^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{1j}\right\}\right]$$

and

$$\tilde{\tilde{w}}_i(\mathcal{N}_2)=\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\hat{e}_j\left[w_{ij}-\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)^\top\hat{\boldsymbol{\Sigma}}^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)w_{ij}\right\}\right].$$

Proof of Lemma 18. *Step 1.* Prove

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{\tilde{w}}_i\left(\mathcal{N}_2\right)\right\}^2=\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2+o_p(1).$$

Similar to *Step 3* in the proof of Lemma 16, we must prove that

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{\tilde{w}}_i\left(\mathcal{N}_2\right)\right\}^2=\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2+o_p(1),\tag{S2.162}$$

as the proofs of the other terms are similar.

Define $\hat{\mathcal{H}}^{(ij)}=\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)^\top\hat{\boldsymbol{\Sigma}}^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)w_{ij}\right\}$ and $\mathcal{H}^{(ij)}=\dot{g}(\boldsymbol{\theta}_0,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}_*^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}_0,\boldsymbol{X}_1\right)w_{1j}\right\}$. Thus,

$$\begin{aligned}&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{\tilde{w}}_i\left(\mathcal{N}_2\right)\right\}^2-\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2\\=&\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\left(w_{ij}-\hat{\mathcal{H}}^{(ij)}\right)\right\}^2\\&-\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{l(\boldsymbol{X}_i)\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\left(w_{ij}-\mathcal{H}^{(ij)}\right)\right\}^2\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\hat{\mathcal{H}}^{(ij)}\right)^2-\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\mathcal{H}^{(ij)}\right)^2\\
&-\frac{2}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\hat{\mathcal{H}}^{(ij)}\right)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j w_{ij}\right)\\
&+\frac{2}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\mathcal{H}^{(ij)}\right)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j w_{ij}\right). \qquad \text{(S2.163)}
\end{aligned}
$$

As an example, we prove that

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\hat{\mathcal{H}}^{(ij)}\right)^2\\
&-\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\hat{e}_j\mathcal{H}^{(ij)}\right)^2=o_p\left(1\right).
\end{aligned}
$$

We can handle the other two terms in the last two lines of (S2.163) by a similar method. Similar to the proof of $D_1'^{(1)}-D_1^{(1)}=o_p\left(\max\left\{1,n_2\delta_n^2\right\}\right)$ in *Step 3* of the proof of Lemma 16, we only illustrate that

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left\{\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\hat{\mathcal{H}}^{(ij)}\right\}^2\\
&-\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left\{\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\mathcal{H}^{(ij)}\right\}^2=o_p\left(1\right). \qquad \text{(S2.164)}
\end{aligned}
$$

To prove (S2.164), we use Lemma 15 again. Note that

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left\{\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\mathcal{H}^{(ij)}\right\}^2\\
=&\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{1j}\right\}^\top\boldsymbol{\Sigma}^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)^\top\\
&\times\boldsymbol{\Sigma}^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{1j}\right\} \qquad \text{(S2.165)}
\end{aligned}
$$

and

$$\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\left\{\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\hat{\mathcal{H}}^{(ij)}\right\}^2 \\
=&\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)w_{ij}\right\}\hat{\boldsymbol{\Sigma}}^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)^{\top} \\
&\times\hat{\boldsymbol{\Sigma}}^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)w_{ij}\right\}. \qquad \text{(S2.166)}
\end{aligned}$$

Under Condition 4, $\left\|\boldsymbol{\Sigma}^{-1}\right\| = O(1)$, $\left\|\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)^{\top}\right\| =$ $O_p(1)$, and according to (S2.154),

$$\begin{aligned}
&\left\|\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{1j}\right\}\right\| \\
=&\left\|E\left\{l(\boldsymbol{X}_2)\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{12}\right\}\right\| + o_p(1) = O_p(1).
\end{aligned}$$

Similar to (S2.134) and (S2.135), we conclude that

$$\begin{aligned}
&\left\|\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)^{\top}\right. \\
&\left.-\frac{1}{n_1}\sum_{i=1}^{n_1} l^2(\boldsymbol{X}_i)\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_i\right)^{\top}\right\| = o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
&\left\|\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_i\right)w_{ij}\right\}\right. \\
&\left.-\frac{1}{n_2}\sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{1j}\right\}\right\| = o_p(1).
\end{aligned}$$

From the proof of (S2.145), we infer that $\left\|\hat{\boldsymbol{\Sigma}}^{-1}-\boldsymbol{\Sigma}^{-1}\right\| = o(1)$. Thus,

combining (S2.165), (S2.166) and Lemma 15, we know that (S2.164) holds. Then we infer that (S2.162) holds, and *Step 1* is finished.

Step 2. Prove

$$\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{\tilde{w}}_i\left(\mathcal{N}_2\right)\right\}^2=\left\{\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\frac{1}{\sqrt{n_2}}\tilde{w}_i^{(0)}\left(\mathcal{N}_2\right)\right\}^2+o_p\left(1\right). \tag{S2.167}$$

The analysis of (S2.167) is similar to (S2.147) and (S2.148). We show that

$$\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_i)l(\boldsymbol{X}_j)\left(\mathcal{H}^{(ij)}-\hat{\mathcal{H}}^{(ij)}\right)=o_p\left(1\right). \tag{S2.168}$$

since the other terms converging to zero can be proved by easier methods. Here,

$$\begin{aligned}
&\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_i)l(\boldsymbol{X}_j)\mathcal{H}^{(ij)}\\
=&\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_i)l(\boldsymbol{X}_j)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}_*^{-1}E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{1j}\right\}\\
=&\frac{1}{n_1}\sum_{i=1}^{n_1}l(\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)^\top\boldsymbol{\Sigma}_*^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)E_{\boldsymbol{X}_j}\left\{\dot{g}\left(\boldsymbol{\theta}^*,\boldsymbol{X}_1\right)w_{1j}\right\}
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{n_1n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_i)l(\boldsymbol{X}_j)\hat{\mathcal{H}}^{(ij)}\\
=&\frac{1}{n_1}\sum_{i_1=1}^{n_1}l(\boldsymbol{X}_{i_1})\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_{i_1}\right)^\top\boldsymbol{\Sigma}_*^{-1}\frac{1}{n_2}\sum_{j=n_1+1}^{n}l(\boldsymbol{X}_j)\frac{1}{n_1}\sum_{i_2=1}^{n_1}\left\{\dot{g}\left(\hat{\boldsymbol{\theta}},\boldsymbol{X}_{i_2}\right)w_{i_2j}\right\}.
\end{aligned}$$

Similar to *Step 1*, we have

$$\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}l(\boldsymbol{X}_i)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_i)\right\|=\left\|E\left\{l(\boldsymbol{X}_1)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)\right\}\right\|+o_p\left(1\right)=o_p\left(1\right),$$

$$\left\| \frac{1}{n_2} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j) E_{\boldsymbol{X}_j} \left\{ \dot{g} \left(\boldsymbol{\theta}^*, \boldsymbol{X}_1 \right) w_{1j} \right\} \right\|$$

$$= \left\| E \left\{ l(\boldsymbol{X}_2) \dot{g} \left(\boldsymbol{\theta}^*, \boldsymbol{X}_1 \right) w_{12} \right\} \right\| + o_p(1) = O_p(1),$$

$$\left\| \frac{1}{n_1} \sum_{i=1}^{n_1} l(\boldsymbol{X}_i) \dot{g} \left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i \right)^\top - \frac{1}{n_1} \sum_{i=1}^{n_1} l(\boldsymbol{X}_i) \dot{g} (\boldsymbol{\theta}^*, \boldsymbol{X}_i) \right\| = o_p(1),$$

$$\left\| \frac{1}{n_2} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j) \frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \dot{g} \left(\hat{\boldsymbol{\theta}}, \boldsymbol{X}_i \right) w_{ij} \right\} \right.$$
$$\left. - \frac{1}{n_2} \sum_{j=n_1+1}^{n} l(\boldsymbol{X}_j) E_{\boldsymbol{X}_j} \left\{ \dot{g} \left(\boldsymbol{\theta}^*, \boldsymbol{X}_1 \right) w_{1j} \right\} \right\| = o_p(1),$$

$\left\| \boldsymbol{\Sigma}_*^{-1} \right\| = O_p(1)$ and $\left\| \hat{\boldsymbol{\Sigma}}^{-1} - \boldsymbol{\Sigma}_*^{-1} \right\| = o_p(1)$. Thus, using Lemma 15, we infer that (S2.168) holds.

Step 2 is finished. From the results of *Step 1* and *Step 2*, we conclude that Lemma 18 holds. The proof is finished.

□

To prove the asymptotic normality of the test statistic under the null hypothesis, we introduce Berry-Esseen Theorem as follows.

Lemma 19. *(Berry-Esseen Theorem) If* $\{X_i, i \geq 1\}$ *are i.i.d. mean-zero random variables with* $E|X_1|^2 = \sigma^2, E|X_1|^3 < \infty$*, then there exists a constant* $C_\kappa > 0$ *such that*

$$\sup_{-\infty < x < \infty} \left| \mathbb{P} \left(\sum_{i=1}^{n} X_i < x \sigma n^{1/2} \right) - \Phi(x) \right| \leq C_\kappa \frac{E|X_1|^3}{\sigma^3 \sqrt{n}}.$$

Now we start to prove **Theorem 1**.

Proof of Theorem 1. To prove that $COST_n$ has a standard normal weak limit, we define

$$T_n = \frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij}}{\sqrt{E_{\mathcal{N}_2} \left\{ \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right)^2 \right\}}},$$

where w'_{ij} is defined in Condition 9. We first show the convergence of T_n and then the asymptotic equivalence between $COST_n$ and T_n.

Under the null hypothesis, according to Lemma 10, we have

$$R(\mathcal{N}_2) := \frac{E_{\mathcal{N}_2} \left(\left| \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right|^3 \right)}{E_{\mathcal{N}_2}^{3/2} \left\{ \log^2 n_1 \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right)^2 \right\}} = O_p(1). \quad \text{(S2.169)}$$

Then according to Lemma 19,

$$\begin{aligned} D(\mathcal{N}_2) &:= \sup_{-\infty<x<\infty} \left| \mathbb{P}\left(T_n < x | \mathcal{N}_2\right) - \Phi(x) \right| \\ &\leq \min\left\{ 1, C_\kappa \frac{R(\mathcal{N}_2) \log^3 n_1}{\sqrt{n_1}} \right\} = O_p\left(\frac{\log^3 n_1}{\sqrt{n_1}} \right). \end{aligned}$$

Because $D(\mathcal{N}_2) \leq 1$, $D(\mathcal{N}_2)$ is uniformly integrable. Then $\lim_{n_1,n_2\to\infty} E\left\{D(\mathcal{N}_2)\right\} = 0$. That is to say,

$$\begin{aligned} &\lim_{n_1,n_2\to\infty} \sup_{-\infty<x<\infty} \left| \mathbb{P}\left(T_n < x\right) - \Phi(x) \right| \\ \leq &\lim_{n_1,n_2\to\infty} E\left\{ \sup_{-\infty<x<\infty} \left| \mathbb{P}\left(T_n < x | \mathcal{N}_2\right) - \Phi(x) \right| \right\} = 0. \end{aligned}$$

So $T_n \xrightarrow{d} \mathbf{N}(0,1)$.

Next, from Lemma 5, we have

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} = \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + o_p(1).$$

From (S2.69), we know that for any random variable sequence $\xi_n = o_p(1)$,

$$\frac{\xi_n}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j}\right)^2\right\}} = o_p(1).$$

Then it follows that

$$\frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}}{\sqrt{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j}\right)^2\right\}}} = T_n + o_p(1).$$

According to case (a) in Lemma 16, we have

$$\frac{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{\hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j}\right)^2\right\}} \xrightarrow{p} 1,$$

where " $\xrightarrow{p}$ " stands for convergence in probability. Using Slutsky Theorem, we have

$$COST_n = \frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}}{\sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{\hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}} \xrightarrow{d} \mathbf{N}(0,1).$$

The proof is finished. □

The following corollary states that V_n, mentioned in the main context, also has a normal weak limit under the null hypothesis in (S2.6). This motivates us to construct our statistic $COST_n$ in the main context.

Corollary 3. *If Conditions 1-9 hold, under the null hypothesis,*

$$V_n = \frac{\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} e_i \tilde{w}_i(\mathcal{N}_2)}{\sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ e_i \tilde{w}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} e_i \tilde{w}_i(\mathcal{N}_2) \right\}^2}} \xrightarrow{d} \mathbf{N}(0,1).$$

Proof of Corollary 3. Recall that under the null hypothesis, $\varepsilon_i = e_i$.

From Lemma 6, we know

$$\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} e_i \tilde{w}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2) = \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_i \varepsilon_j w'_{ij} + o_p(1).$$

From (S2.109), we know under the null hypothesis with $\delta_n = 0$,

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \{\varepsilon_i \tilde{w}_i(\mathcal{N}_2)\}^2 = E_{\mathcal{N}_2} \left\{ \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right)^2 \right\} + o_p(1).$$

From (S2.123), we can derive that under the null hypothesis with $\delta_n = 0$,

$$\left\{ \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2) \right\}^2 = o_p(1).$$

So

$$\begin{aligned} &\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \varepsilon_i \tilde{w}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2) \right\}^2 \\ =& E_{\mathcal{N}_2} \left\{ \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right)^2 \right\} + o_p(1). \end{aligned}$$

Taking advantage of (S2.69), we derive that

$$\frac{\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \varepsilon_i \tilde{w}_i(\mathcal{N}_2)}{E_{\mathcal{N}_2} \left\{ \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right)^2 \right\}} = \frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j}}{E_{\mathcal{N}_2} \left\{ \left(\frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} \right)^2 \right\}} + o_p(1) \tag{S2.170}$$

and

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\varepsilon_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\varepsilon_i\tilde{w}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}-1=o_p(1). \tag{S2.171}$$

According to the proof of **Theorem 1**, we have

$$T_n=\frac{\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij}}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}}\xrightarrow{d}\mathbf{N}(0,1). \tag{S2.172}$$

Thus, using (S2.170), (S2.171), (S2.172) and Slutsky Theorem, since $e_i=\varepsilon_i$, we conclude that

$$V_n=\frac{\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}e_i\tilde{w}_i(\mathcal{N}_2)}{\sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{e_i\tilde{w}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}e_i\tilde{w}_i(\mathcal{N}_2)\right\}^2}}\xrightarrow{d}\mathbf{N}(0,1).$$

The proof is finished. □

Proof of Theorem 2. **Deal with case (a).** From Lemma 11, we have

$$\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij}=E\left\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}+o_p(1).$$

Incorporating Lemma 17 with Lemma 18, we know that

$$\begin{aligned}
&\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2\\
=&E\left[\left\{\varepsilon_1^2+l^2(\boldsymbol{X}_1)\right\}E^2_{\boldsymbol{X}_1}\left\{l(\boldsymbol{X}_2)w_{12}\right\}\right]-E^2\left\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\right\}\\
&+E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\}^\top\boldsymbol{\Sigma}_*^{-1}\left(\boldsymbol{\Sigma}_\varepsilon+\boldsymbol{\Sigma}_l\right)\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\}\\
&-2E\left[\left\{\varepsilon_1^2+l^2(\boldsymbol{X}_1)\right\}l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right]^\top\boldsymbol{\Sigma}_*^{-1}E\left\{l(\boldsymbol{X}_2)\dot{g}(\boldsymbol{\theta}^*,\boldsymbol{X}_1)w_{12}\right\}
\end{aligned}$$

$$+ o_p(1)$$

$$=: V_{(0)} + o_p(1).$$

Thus, using Slutsky Theorem, we obtain that

$$\begin{aligned}\frac{COST_n}{\sqrt{n_1}} =& \frac{\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij}}{\sqrt{\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}} \\ \xrightarrow{p}& \frac{E\left[\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)\}\, w_{12}\right]}{\sqrt{V_{(0)}}}.\end{aligned}$$

Case (a) is proved.

Deal with case (b). From Lemma 12,

$$\begin{aligned}&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij} \\ =&\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + \frac{\sqrt{n_2}}{\sqrt{n_1 n}}\sum_{i=1}^{n_1}\varepsilon_i H(\boldsymbol{X}_i) \\ &+ \frac{\sqrt{n_1}}{\sqrt{n_2 n}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) + \frac{\sqrt{n_1 n_2}}{n}\mu + o_p(1),\end{aligned}$$

where

$$\begin{aligned}H(\boldsymbol{X}_1) =& E_{\boldsymbol{X}_1}\{l(\boldsymbol{X}_2)w_{12}\} - E_{\boldsymbol{X}_1}\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_2)w_{12}\}^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\} \\ &- \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\} \\ &+ \dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}^*_{cov}\boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\},\end{aligned}$$

and

$$\mu = E\{l(\boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\} - 2E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_2)w_{12}\}^\top \boldsymbol{\Sigma}^{-1} E\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1)l(\boldsymbol{X}_1)\}$$

$$+ E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^*_{cov} \boldsymbol{\Sigma}^{-1} E\left\{\dot{g}(\boldsymbol{\theta}_0, \boldsymbol{X}_1) l(\boldsymbol{X}_1)\right\}.$$

From Lemma 16, we know that

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} \xrightarrow{p} 1.$$

According to Lemma 13,

$$\frac{E_{\mathcal{N}_2}\left\{\left|\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right|^3\right\}}{E^{3/2}_{\mathcal{N}_2}\left[\log^2 n_1\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} = O_p(1).$$

Similar to the proof of **Theorem 1**, using Berry-Esseen Theorem in Lemma 19, we have

$$\frac{\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + \frac{\sqrt{n_2}}{\sqrt{n_1 n}}\sum_{i=1}^{n_1}\varepsilon_i H(\boldsymbol{X}_i)}{\sqrt{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \frac{\sqrt{n_2}}{\sqrt{n}}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}} \xrightarrow{d} \mathbf{N}(0,1).$$

From case (a) in Lemma 12 and case (a) in Lemma 14, we have

$$\frac{\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij} - \frac{\sqrt{n_1}}{\sqrt{n_2 n}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) - \frac{\sqrt{n_1 n_2}}{n}\mu}{\sqrt{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \frac{\sqrt{n_2}}{\sqrt{n}}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}} \xrightarrow{d} \mathbf{N}(0,1). \tag{S2.173}$$

According to case (b) in Lemma 16, we have

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \frac{\sqrt{n_2}}{\sqrt{n}}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} \xrightarrow{p} 1.$$

Then it follows that

$$\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij} \times \left[\frac{\sqrt{E_{\mathcal{N}_2}\left[\left\{ \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} + \frac{\sqrt{n_2}}{\sqrt{n}} \varepsilon_1 H(\boldsymbol{X}_1) \right\}^2 \right]}}{\sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) \right\}^2}} - 1 \right] \xrightarrow{p} 0, \tag{S2.174}$$

Dividing both sides of (S2.174) by

$$\sqrt{E_{\mathcal{N}_2}\left[\left\{ \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} + \frac{\sqrt{n_2}}{\sqrt{n}} \varepsilon_1 H(\boldsymbol{X}_1) \right\}^2 \right]},$$

we have

$$\frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}}{\sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) \right\}^2}} - \frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}}{\sqrt{E_{\mathcal{N}_2}\left[\left\{ \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} + \frac{\sqrt{n_2}}{\sqrt{n}} \varepsilon_1 H(\boldsymbol{X}_1) \right\}^2 \right]}} \xrightarrow{p} 0. \tag{S2.175}$$

Incorporating (S2.173) with (S2.175) and using Slutsky theroem, we derive

$$\frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}}{\sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) \right\}^2}} - \frac{\frac{\sqrt{n_1}}{\sqrt{n_2 n}} \sum_{j=n_1+1}^{n} \varepsilon_j H(\boldsymbol{X}_j) + \frac{\sqrt{n_1 n_2}}{n} \mu}{\sqrt{E_{\mathcal{N}_2}\left[\left\{ \frac{1}{\sqrt{n_2}} \sum_{j=n_1+1}^{n} \varepsilon_1 \varepsilon_j w'_{1j} + \frac{\sqrt{n_2}}{\sqrt{n}} \varepsilon_1 H(\boldsymbol{X}_1) \right\}^2 \right]}} \xrightarrow{d} \mathbf{N}(0,1).$$

Under Condition 10, $\frac{E\left[\{\varepsilon_1 H(\boldsymbol{X}_1)\}^4\right]}{E^2\left[\{\varepsilon_1 H(\boldsymbol{X}_1)\}^2\right]} = O(1)$, hence, according to Berry-Esseen Theorem in Lemma 19, $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) \xrightarrow{d} \mathbf{N}\left(0, E\left[\{\varepsilon_1 H(\boldsymbol{X}_1)\}^2\right]\right)$.

Case (b) is proved.

Deal with case (c). According to the definition of $COST_n$ in the main context,

$$\frac{n^\alpha COST_n}{\sqrt{n_1}} = \frac{\frac{n^\alpha}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij}}{\sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}}. \tag{S2.176}$$

For the numerator on the right-hand side in equation (S2.176), from Lemma 12 case (b), we know that when $n^{-2\alpha}p^3\log n \to 0$, it holds that

$$\begin{aligned}
&\frac{n^\alpha}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij} \\
=&\frac{n^\alpha}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + \frac{\sqrt{n_2}}{n_1}\sum_{i=1}^{n_1}\varepsilon_i H(\boldsymbol{X}_i) + \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) \\
&+ \frac{\sqrt{n_2}}{n^\alpha}\mu + o_p\left(\max\left\{1, \sqrt{n_2/n_1}, \sqrt{n_2}n^{-\alpha}\right\}\right).
\end{aligned}$$

When $n^{-\alpha}\sqrt{n_1} \to \infty$, $\frac{n^\alpha}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} = O_p\left(n^\alpha/\sqrt{n_1}\right) = o_p(1)$. Thus,

$$\begin{aligned}
&\frac{n^\alpha}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij} \\
=&\frac{\sqrt{n_2}}{n_1}\sum_{i=1}^{n_1}\varepsilon_i H(\boldsymbol{X}_i) + \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) \\
&+ \frac{\sqrt{n_2}}{n^\alpha}\mu + o_p\left(\max\left\{1, \sqrt{n_2/n_1}, \sqrt{n_2}n^{-\alpha}\right\}\right).
\end{aligned} \tag{S2.177}$$

For the denominator in equation (S2.176), according to case (b) in Lemma 16, we have

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left\{\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)-\frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right\}^2}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+n^{-\alpha}\sqrt{n_2}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}-1=o_p(1).$$

Thus, we first analyse

$$\frac{\frac{n^\alpha}{n_1\sqrt{n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij}}{\sqrt{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+n^{-\alpha}\sqrt{n_2}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}}=:V'_n$$

and then we use the Slutsky theorem to get the final result. To this end, we consider three cases: $n^{-\alpha}\sqrt{n_2}\to\infty$, $n^{-\alpha}\sqrt{n_2}=K$ and $n^{-\alpha}\sqrt{n_2}=o(1)$.

Prove case (c.1). If $n^{-\alpha}\sqrt{n_2}\to\infty$ as $n\to\infty$, then from (S2.93) it holds that

$$\begin{aligned}&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+n^{-\alpha}\sqrt{n_2}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]\\ =&E\left[\left\{n^{-\alpha}\sqrt{n_2}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]+o_p\left(n^{-2\alpha}n_2\right).\end{aligned}$$

Thus, the convergence rate of the denominator of V'_n is $O_p\left(n^{-\alpha}\sqrt{n_2}\right)$, and

$$\frac{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}+n^{-\alpha}\sqrt{n_2}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}{E\left[\left\{n^{-\alpha}\sqrt{n_2}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}-1=o_p(1).$$

Note that $\frac{\sqrt{n_2}}{n_1}\sum_{i=1}^{n_1}\varepsilon_i H(\boldsymbol{X}_i)=O_p\left(\sqrt{n_2/n_1}\right)=o_p\left(n^{-\alpha}\sqrt{n_2}\right)$ and $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j)=O_p(1)=o_p\left(n^{-\alpha}\sqrt{n_2}\right)$. Thus, from (S2.177) we

know

$$V'_n = \frac{n^{-\alpha}\sqrt{n_2}\mu}{\sqrt{E\left[\left\{n^{-\alpha}\sqrt{n_2}\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}} + o_p(1) = \frac{\mu}{\sqrt{E\left[\left\{\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}} + o_p(1).$$

Using Slutsky theorem, we have

$$\frac{n^{\alpha} COST_n}{\sqrt{n_1}} \xrightarrow{p} \frac{\mu}{\sqrt{E\left[\left\{\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}}.$$

Prove case (c.2). If $n^{-\alpha}\sqrt{n_2} \to K$ as $n \to \infty$, where K is a positive constant, then $\frac{\sqrt{n_2}}{n_1}\sum_{i=1}^{n_1}\varepsilon_i H(\boldsymbol{X}_i) = O_p\left(\sqrt{n_2/n_1}\right) = o_p(1)$, $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) = O_p(1)$ and $\frac{\sqrt{n_2}}{n^{\alpha}}\mu \to K\mu = O(1)$. Thus,

$$V'_n = \frac{K\mu + \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j)}{\sqrt{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + K\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}} + o_p(1).$$

and it leads to

$$\frac{n^{\alpha} COST_n}{\sqrt{n_1}} = \frac{K\mu + \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j)}{\sqrt{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + K\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}} + o_p(1),$$

where $K\mu + \frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) \xrightarrow{d} \mathbf{N}\left(K\mu, E\left[\left\{\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]\right)$.

Prove case (c.3). If $n^{-\alpha}\sqrt{n_2} = o(1)$, we have $\frac{\sqrt{n_2}}{n_1}\sum_{i=1}^{n_1}\varepsilon_i H(\boldsymbol{X}_i) = O_p\left(\sqrt{n_2/n_1}\right) = o_p(1)$, $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) = O_p(1)$ and $\frac{\sqrt{n_2}}{n^{\alpha}}\mu = o(1)$. This is similar to the other two cases, we have

$$\frac{n^{\alpha} COST_n}{\sqrt{n_1}} = \frac{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j)}{\sqrt{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right\}^2\right]}} + o_p(1),$$

where $\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_j H(\boldsymbol{X}_j) \xrightarrow{d} \mathbf{N}\left(0, E\left[\{\varepsilon_1 H(\boldsymbol{X}_1)\}^2\right]\right)$.

Case (c) is proved.

Deal with case (d). From case (c) in Lemma 12, if $\delta_n = n^{-\alpha}$, $\alpha > \frac{1}{2}$, $p^3 \log n_1/n_1 \to 0$ and $p^3 \log n_2/n_2 \to 0$, then

$$\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij} = \frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\varepsilon_i\varepsilon_j w'_{ij} + o_p(1).$$

From case (b) in Lemma 16, we have

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{e}_i\tilde{\tilde{w}}_i(\mathcal{N}_2)\right]^2}{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]} \xrightarrow{p} 0. \tag{S2.178}$$

Note that $\sqrt{n_2}\delta_n = o(1)$, so

$$\begin{aligned}&E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]\\ =&E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\} + o_p(1),\end{aligned}$$

and from (S2.69), we conclude that

$$\frac{E_{\mathcal{N}_2}\left[\left\{\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j} + \sqrt{n_2}\delta_n\varepsilon_1 H(\boldsymbol{X}_1)\right\}^2\right]}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}} \xrightarrow{p} 1, \tag{S2.179}$$

Similar to the proof of **Theorem 1**, we have

$$\frac{\frac{1}{\sqrt{n_1 n_2}}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n}\hat{e}_i\hat{e}_j w_{ij}}{E_{\mathcal{N}_2}\left\{\left(\frac{1}{\sqrt{n_2}}\sum_{j=n_1+1}^{n}\varepsilon_1\varepsilon_j w'_{1j}\right)^2\right\}} \xrightarrow{p} \mathbf{N}(0,1). \tag{S2.180}$$

Using (S2.178), (S2.179), (S2.180) and Slutsky theorem, we arrive at the conclusion

$$COST_n = \frac{\frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} \hat{e}_i \hat{e}_j w_{ij}}{\sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} \left\{ \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) - \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{e}_i \tilde{\tilde{w}}_i(\mathcal{N}_2) \right\}^2}} \xrightarrow{d} \mathbf{N}(0,1).$$

Case (d) is proved. That is the end of the proof of **Theorem 2**.

□

Bibliography

Hall, P. G. and Heyde, C. C. (1980). Martingale limit theory and its application. *Journal of the American Statistical Association*, 78(381).

Tan, F. and Zhu, L. (2022). Integrated conditional moment test and beyond: when the number of covariates is divergent. *Biometrika*, **109**:103–122.

Zhu, L. and Fang, K. (1996). Asymptotics for kernel estimate of sliced inverse regression. *The Annals of Statistics*, 24(3).