# Supplementary Material to "Optimal linear discriminant analysis for high-dimensional functional data"

## 1 Notations

First we recall the basic notations used throughout the paper. For every $j \le p_n$, consider the diagonal matrices or structures

$$\Lambda_j = \mathrm{diag}\{\omega_{j1}, \omega_{j2}, \dots\}, \quad \Lambda_j^{(1)} = \mathrm{diag}\{\omega_{j1}, \dots, \omega_{js_n}\}, \quad \Lambda_j^{(2)} = \mathrm{diag}\{\omega_{j,s_n+1}, \omega_{j,s_n+2}, \dots\},$$

$$\hat{\Lambda}_j = \mathrm{diag}\{\hat{\omega}_{j1}, \hat{\omega}_{j2}, \dots\}, \quad \hat{\Lambda}_j^{(1)} = \mathrm{diag}\{\hat{\omega}_{j1}, \dots, \hat{\omega}_{js_n}\}, \quad \hat{\Lambda}_j^{(2)} = \mathrm{diag}\{\hat{\omega}_{j,s_n+1}, \hat{\omega}_{j,s_n+2}, \dots\},$$

we then denote several block matrices or structures as

$$\Lambda = \mathrm{diag}\{\Lambda_j : j \le p_n\}, \quad \Lambda^{(1)} = \mathrm{diag}\{\Lambda_j^{(1)} : j \le p_n\}, \quad \Lambda^{(2)} = \mathrm{diag}\{\Lambda_j^{(2)} : j \le p_n\},$$

$$\Lambda_T = \mathrm{diag}\{\Lambda_j : j \in T\}, \quad \Lambda_T^{(1)} = \mathrm{diag}\{\Lambda_j^{(1)} : j \in T\}, \quad \Lambda_T^{(2)} = \mathrm{diag}\{\Lambda_j^{(2)} : j \in T\},$$

$$\hat{\Lambda} = \mathrm{diag}\{\hat{\Lambda}_j : j \le p_n\}, \quad \hat{\Lambda}^{(1)} = \mathrm{diag}\{\hat{\Lambda}_j^{(1)} : j \le p_n\}, \quad \hat{\Lambda}^{(2)} = \mathrm{diag}\{\hat{\Lambda}_j^{(2)} : j \le p_n\},$$

$$\hat{\Lambda}_T = \mathrm{diag}\{\hat{\Lambda}_j : j \in T\}, \quad \hat{\Lambda}_T^{(1)} = \mathrm{diag}\{\hat{\Lambda}_j^{(1)} : j \in T\}, \quad \hat{\Lambda}_T^{(2)} = \mathrm{diag}\{\hat{\Lambda}_j^{(2)} : j \in T\}.$$

Similar to the constructions of $\xi^{(1)}$ and $\xi_T^{(1)}$, we let $\xi^{(2)} = (\tilde{\xi}_1^{(2)\prime}, \dots, \tilde{\xi}_{p_n}^{(2)\prime})'$ with sub-vectors $\tilde{\xi}_j^{(2)} = (\xi_{j,s_n+1}, \xi_{j,s_n+2}, \dots)'$, and $\xi_T^{(2)}$ as stacking $\{\tilde{\xi}_j^{(2)} : j \in T\}$ in a column. Given index sets $T$ and $N$, we define several covariance matrices and structures as

$$\Sigma_{TT}^{(1)} = \mathrm{var}(\xi_T^{(1)}), \quad \Sigma_{NN}^{(1)} = \mathrm{var}(\xi_N^{(1)}), \quad \Sigma_{TN}^{(1)} = \mathrm{cov}(\xi_T^{(1)}, \xi_N^{(1)}), \quad \Sigma_{NT}^{(1)} = \mathrm{cov}(\xi_N^{(1)}, \xi_T^{(1)}),$$

$$\Sigma_{TT}^{(2)} = \mathrm{var}(\xi_T^{(2)}), \quad \Sigma_{NN}^{(2)} = \mathrm{var}(\xi_N^{(2)}), \quad \Sigma_{TN}^{(2)} = \mathrm{cov}(\xi_T^{(2)}, \xi_N^{(2)}), \quad \Sigma_{NT}^{(2)} = \mathrm{cov}(\xi_N^{(2)}, \xi_T^{(2)}),$$

$$\Sigma_{TT}^{(1,2)} = \mathrm{cov}(\xi_T^{(1)}, \xi_T^{(2)}), \quad \Sigma_{NN}^{(1,2)} = \mathrm{cov}(\xi_N^{(1)}, \xi_N^{(2)}), \quad \Sigma_{TN}^{(1,2)} = \mathrm{cov}(\xi_T^{(1)}, \xi_N^{(2)}),$$

$$\Sigma_{NT}^{(1,2)} = \mathrm{cov}(\xi_N^{(1)}, \xi_T^{(2)}), \quad \Sigma_{TT}^{(2,1)} = \mathrm{cov}(\xi_T^{(2)}, \xi_T^{(1)}), \quad \Sigma_{NN}^{(2,1)} = \mathrm{cov}(\xi_N^{(2)}, \xi_N^{(1)}),$$

$$\Sigma_{TN}^{(2,1)} = \mathrm{cov}(\xi_T^{(2)}, \xi_N^{(1)}), \quad \Sigma_{NT}^{(2,1)} = \mathrm{cov}(\xi_N^{(2)}, \xi_T^{(1)}).$$

Similar to the constructions of the vectors $\xi_T^{(1)}$, $\mu_{1,T}^{(1)}$, $\mu_{2,T}^{(1)}$, and $\nu_T^{(1)}$, we define $\xi_{i,T}^{(1)}$, $\hat{\mu}_{1,T}^{(1)}$, $\hat{\mu}_{2,T}^{(1)}$, and $\hat{\nu}_T^{(1)}$ as restricting the vectors $\xi_i^{(1)}$, $\hat{\mu}_1^{(1)}$, $\hat{\mu}_2^{(1)}$, and $\hat{\nu}^{(1)}$ to the discriminant set $T$. Given index sets $T$ and $N$, we define several sample covariance matrices as

$$S^{(1)} = \{(n_1 - 1)S_1^{(1)} + (n_2 - 1)S_2^{(1)}\}/(n-2),$$

$$S_{TT}^{(1)} = \{(n_1 - 1)S_{1,TT}^{(1)} + (n_2 - 1)S_{2,TT}^{(1)}\}/(n-2),$$

$$S_{NT}^{(1)} = \{(n_1 - 1)S_{1,NT}^{(1)} + (n_2 - 1)S_{2,NT}^{(1)}\}/(n-2),$$

where

$$S_1^{(1)} = \sum_{i \in H_1} (\xi_i^{(1)} - \hat{\mu}_1^{(1)})(\xi_i^{(1)} - \hat{\mu}_1^{(1)})'/(n_1 - 1),$$

$$S_2^{(1)} = \sum_{i \in H_2} (\xi_i^{(1)} - \hat{\mu}_2^{(1)})(\xi_i^{(1)} - \hat{\mu}_2^{(1)})'/(n_2 - 1),$$

$$S_{1,TT}^{(1)} = \sum_{i \in H_1} (\xi_{i,T}^{(1)} - \hat{\mu}_{1,T}^{(1)})(\xi_{i,T}^{(1)} - \hat{\mu}_{1,T}^{(1)})'/(n_1 - 1),$$

$$S_{2,TT}^{(1)} = \sum_{i \in H_2} (\xi_{i,T}^{(1)} - \hat{\mu}_{2,T}^{(1)})(\xi_{i,T}^{(1)} - \hat{\mu}_{2,T}^{(1)})'/(n_2 - 1),$$

$$S_{1,NT}^{(1)} = \sum_{i \in H_1} (\xi_{i,N}^{(1)} - \hat{\mu}_{1,N}^{(1)})(\xi_{i,T}^{(1)} - \hat{\mu}_{1,T}^{(1)})'/(n_1 - 1),$$

$$S_{2,NT}^{(1)} = \sum_{i \in H_2} (\xi_{i,N}^{(1)} - \hat{\mu}_{2,N}^{(1)})(\xi_{i,T}^{(1)} - \hat{\mu}_{2,T}^{(1)})'/(n_2 - 1).$$

Similar to the definitions of $\mu_1^{(1)}$, $\mu_2^{(1)}$, $\nu^{(1)}$, $\mu_{1,T}^{(1)}$, $\mu_{2,T}^{(1)}$, and $\nu_T^{(1)}$, we denote for any $\ell = 1, 2$,

$$\mu_\ell^{(2)} = E(\xi^{(2)} | Y = \ell) = (\tilde{\mu}_{\ell 1}^{(2)\prime}, \dots, \tilde{\mu}_{\ell p_n}^{(2)\prime})',$$

$$\tilde{\mu}_{\ell j}^{(2)} = E(\tilde{\xi}_j^{(2)} | Y = \ell) = (\mu_{\ell j, s_n+1}, \mu_{\ell j, s_n+2}, \dots)' \in \mathbb{R}^\infty, \quad j = 1, \dots, p_n,$$

$$\mu_{\ell,T}^{(2)} : \quad \text{formed by stacking} \quad \{\tilde{\mu}_{\ell j}^{(2)} : j \in T\} \quad \text{in a column},$$

$$\nu^{(2)} = \mu_2^{(2)} - \mu_1^{(2)}, \quad \nu_T^{(2)} = \mu_{2,T}^{(2)} - \mu_{1,T}^{(2)}.$$

Similar to the constructions of $\beta^{(1)}$ and $\beta_T^{(1)}$, we denote $\beta^{*(1)}$, $\beta_T^{*(1)}$, $\beta^{*(2)}$, and $\beta_T^{*(2)}$ as

$$\beta^{*(1)} = (\beta_1^{*(1)'}, \dots, \beta_{p_n}^{*(1)'})' \text{ with each } \beta_j^{*(1)} = (\beta_{j1}^*, \dots, \beta_{js_n}^*)',$$

$$\beta_T^{*(1)}: \quad \text{formed by stacking } \{\beta_j^{*(1)} : j \in T\} \text{ in a column},$$

$$\beta^{*(2)} = (\beta_1^{*(2)'}, \dots, \beta_{p_n}^{*(2)'})' \text{ with each } \beta_j^{*(2)} = (\beta_{j,s_n+1}^*, \beta_{j,s_n+2}^*, \dots)',$$

$$\beta_T^{*(2)}: \quad \text{formed by stacking } \{\beta_j^{*(2)} : j \in T\} \text{ in a column}.$$

In the next section, we present the proofs of the main results, Theorems 1-2 and Corollary 1.

## 2 Proofs of Theorems 1-2 and Corollary 1

*Proof of Theorem 1*: Under conditions (A1) and (A2), property (i) holds directly from Lemma 1. To show property (ii), first note that

$$\begin{aligned}\Delta =&(\beta_{T^*}^{*'}\Sigma_{T^*T^*}\beta_{T^*}^*)^{1/2} = \{(\Lambda_{T^*}^{1/2}\beta_{T^*}^*)'(\Lambda_{T^*}^{\dagger 1/2}\Sigma_{T^*T^*}\Lambda_{T^*}^{\dagger 1/2})(\Lambda_{T^*}^{1/2}\beta_{T^*}^*)\}^{1/2} \\ \geq& c_1^{1/2}\|\Lambda_{T^*}^{1/2}\beta_{T^*}^*\|_2 = c_1^{1/2}(\sum_{j\in T^*}\sum_{k=1}^{\infty}\omega_{jk}\beta_{jk}^{*2})^{1/2}.\end{aligned}$$

Together with condition (A3), it can be seen that

$$\Delta \to \infty, \quad \text{as} \quad n \to \infty. \tag{1}$$

Hence, property (ii) holds from (6) in the main paper and (1). To show property (iii), first note that

$$\Delta^{(1)} = \{1 + o(r_n^{-1}) + o(r_n^{-1/2}\alpha_n^{1/2})\}\Delta \to \infty, \tag{2}$$

by Lemma 1 and (1). Moreover, by definition, it is not hard to verify that

$$R(\beta^*)/R^{\circ}(\beta^{(1)}) = (\pi_1 + \pi_2\Omega_1)(\pi_1 + \pi_2\Omega_2)^{-1}\Omega_3, \tag{3}$$

where

$$\Omega_1 = \Phi(-\Delta/2 + \log(\pi_1/\pi_2)/\Delta)/\Phi(-\Delta/2 + \log(\pi_2/\pi_1)/\Delta),$$

$$\Omega_2 = \Phi(-\Delta^{(1)}/2 + \log(\pi_1/\pi_2)/\Delta^{(1)})/\Phi(-\Delta^{(1)}/2 + \log(\pi_2/\pi_1)/\Delta^{(1)}),$$

$$\Omega_3 = \Phi(-\Delta/2 + \log(\pi_2/\pi_1)/\Delta)/\Phi(-\Delta^{(1)}/2 + \log(\pi_2/\pi_1)/\Delta^{(1)}).$$

For the term $\Omega_1$, it can be rewritten as

$$\Omega_1 = \Phi(-\varrho_n^{1/2}(1+\vartheta_n))/\Phi(-\varrho_n^{1/2}), \tag{4}$$

where $\varrho_n = \{\Delta/2 - \log(\pi_2/\pi_1)/\Delta\}^2$ and $\vartheta_n = 4\log(\pi_2/\pi_1)/\{\Delta^2 - 2\log(\pi_2/\pi_1)\}$. Since $\varrho_n \to \infty$ and $\varrho_n\vartheta_n \to \log(\pi_2/\pi_1)$ under (1), we immediately conclude that

$$\Omega_1 \to \pi_2/\pi_1, \tag{5}$$

by applying Lemma 1 of Shao et al. (2011) to (4). Similar argument leads to

$$\Omega_2 \to \pi_2/\pi_1. \tag{6}$$

For the term $\Omega_3$, it can be expressed as

$$\Omega_3 = \Phi(-\tilde{\varrho}_n^{1/2}(1+\tilde{\vartheta}_n))/\Phi(-\tilde{\varrho}_n^{1/2}), \tag{7}$$

where $\tilde{\varrho}_n = \{\Delta^{(1)}/2 - \log(\pi_2/\pi_1)/\Delta^{(1)}\}^2$ and $\tilde{\vartheta}_n = [\{\Delta\Delta^{(1)} + 2\log(\pi_2/\pi_1)\}(\Delta - \Delta^{(1)})]/\{\Delta\Delta^{(1)2} - 2\log(\pi_2/\pi_1)\Delta\}$. Based on (2) and (A3), one can show that

$$\tilde{\varrho}_n \to \infty, \quad \tilde{\varrho}_n\tilde{\vartheta}_n \to 0.$$

Together with (7) and Lemma 1 of Shao et al. (2011), it can be concluded that

$$\Omega_3 \to 1.$$

Together with (3), (5) and (6), we have $R(\beta^*)/R^\circ(\beta^{(1)}) \to 1$, which completes the proof. □

**Remark**: Although not part of the proof, it is important to justify that the ideal classifier in (3) of the main article is really the optimal rule. By definition, we have

$$\xi|Y=1 \sim N(\mu_1, \Sigma), \qquad \xi|Y=2 \sim N(\mu_2, \Sigma),$$

which implies

$$\Sigma^{\dagger 1/2}\xi|Y=1 \sim N(\Sigma^{\dagger 1/2}\mu_1, I), \qquad \Sigma^{\dagger 1/2}\xi|Y=2 \sim N(\Sigma^{\dagger 1/2}\mu_2, I).$$

Therefore, the conditional density functions of $z = \Sigma^{\dagger 1/2}\xi$ take the form:

$$f_z(z|Y=i) \propto \exp\{-2^{-1}(z-\Sigma^{\dagger 1/2}\mu_i)'(z-\Sigma^{\dagger 1/2}\mu_i)\}, \quad \text{for} \quad i=1,2.$$

By change of variables, the conditional density functions of $\xi = \Sigma^{1/2}z$ take the form:

$$f_\xi(\xi|Y=i) \propto \exp\{-2^{-1}(\xi-\mu_i)'\Sigma^{\dagger}(\xi-\mu_i)\}, \quad \text{for} \quad i=1,2.$$

Since the optimal rule is such that we assign $\xi$ to the group labeled by $Y=2$ provided that

$$\frac{f_\xi(\xi|Y=1)}{f_\xi(\xi|Y=2)} < \frac{\pi_2}{\pi_1},$$

it can be deduced that the ideal classifier in (3) preserves the optimality of the rule. □

*Proof of Theorem 2*: First of all, it follows from Lemma 11 and the definition of $\hat{v}$ in (16) of the main paper that there exists a universal constant $c_3 > 0$ such that

$$\begin{aligned} & P\{\text{sgn}(\hat{v}) = \text{sgn}(\beta^{(1)})\} \\ \geq & 1 - c_3[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)], \end{aligned} \tag{8}$$

which justifies property (ii). In addition, Lemma 11 also implies that there exists a universal constant $c_4 > 0$ such that

$$\begin{aligned} & P(\hat{v}_T = \tilde{v}_T) \\ \geq & 1 - c_4[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)], \end{aligned} \tag{9}$$

where $\hat{\upsilon}_T$ is defined in (16) of the main paper and

$$\tilde{\upsilon}_T = \{n_1 n_2 n^{-1}(n-2)^{-1}\}\{1+\lambda_n \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})\}\big[1+ \{n_1 n_2 n^{-1}(n-2)^{-1}\}\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\big]^{-1} S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \lambda_n S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}).$$

To prove property (i), based on (8), (9) and the Karush-Kuhn-Tucker conditions, it is sufficient to show that there exist positive constants $c_5, c_6 > 0$ such that

$$P\big[\{n_1 n_2 n^{-1}(n-2)^{-1}\}\hat{\nu}_T^{(1)} - \{S_{TT}^{(1)} + n_1 n_2 n^{-1}(n-2)^{-1}\hat{\nu}_T^{(1)}\hat{\nu}_T^{(1)'}\}\tilde{\upsilon}_T = \lambda_n \hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\tilde{\upsilon}_T)\big] \geq 1 - c_5[\{(p_n-q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)], \tag{10}$$

and

$$P\big(\big\|\hat{\Lambda}_N^{(1)-1/2}\big[\{n_1 n_2 n^{-1}(n-2)^{-1}\}\hat{\nu}_N^{(1)} - \{S_{NT}^{(1)} + n_1 n_2 n^{-1}(n-2)^{-1}\hat{\nu}_N^{(1)}\hat{\nu}_T^{(1)'}\} \cdot \tilde{\upsilon}_T\big]\big\|_\infty \leq \lambda_n\big) \geq 1 - c_6[\{(p_n-q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]. \tag{11}$$

Note that the random quantity $S_{NT}^{(1)}$ can be expressed as $S_{NT}^{(1)} = \{(n_1-1)S_{1,NT}^{(1)} + (n_2-1)S_{2,NT}^{(1)}\}/(n-2)$, where $S_{1,NT}^{(1)} = \sum_{i\in H_1}(\xi_{i,N}^{(1)} - \hat{\mu}_{1,N}^{(1)})(\xi_{i,T}^{(1)} - \hat{\mu}_{1,T}^{(1)})'/(n_1-1)$ and $S_{2,NT}^{(1)} = \sum_{i\in H_2}(\xi_{i,N}^{(1)} - \hat{\mu}_{2,N}^{(1)})(\xi_{i,T}^{(1)} - \hat{\mu}_{2,T}^{(1)})'/(n_2-1)$. Since $\tilde{\upsilon}_T$ is the solution to the convex optimization problem specified in Lemma 2, the first order condition together with Lemma 11 yields (10) immediately. To show (11), we first note that

$$\big\|\hat{\Lambda}_N^{(1)-1/2}\big[\{n_1 n_2 n^{-1}(n-2)^{-1}\}\hat{\nu}_N^{(1)} - \{S_{NT}^{(1)} + n_1 n_2 n^{-1}(n-2)^{-1} \times \hat{\nu}_N^{(1)}\hat{\nu}_T^{(1)'}\}\tilde{\upsilon}_T\big]\big\|_\infty \leq \big(1 + \|\hat{\Lambda}_N^{(1)-1/2}\Lambda_N^{(1)1/2} - I_{(p_n-q_n)s_n}\|_{\max}\big)\cdot\|\Psi\|_\infty, \tag{12}$$

where $\Psi = \Lambda_N^{(1)-1/2}\big[\{S_{NT}^{(1)} + n_1 n_2 n^{-1}(n-2)^{-1}\hat{\nu}_N^{(1)}\hat{\nu}_T^{(1)'}\}\tilde{\upsilon}_T - \{n_1 n_2 n^{-1}(n-2)^{-1}\}\hat{\nu}_N^{(1)}\big]$. By definition, conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\begin{aligned}&(n-2)\Lambda^{(1)-1/2}S^{(1)}\Lambda^{(1)-1/2}|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\\ \sim&\text{Wishart}(n-2|\Lambda^{(1)-1/2}\Sigma^{(1)}\Lambda^{(1)-1/2}).\end{aligned} \tag{13}$$

where the set $\mathcal{M}_n = \{\pi_1/2 \le n_1/n \le 3\pi_1/2\} \cap \{\pi_2/2 \le n_2/n \le 3\pi_2/2\}$ is defined in Lemma 3. Moreover, conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$(n-2)\Lambda^{(1)-1/2}S^{(1)}\Lambda^{(1)-1/2} \quad \perp \quad \hat{\nu}^{(1)},$$

where the symbol $\perp$ means independent of. Together with (13), it can be concluded that there exists a collection $\{Z_l\}_{l=1}^{n-2}$ of $n-2$ random vectors in $\mathbb{R}^{p_n s_n}$ satisfying (14) to (16) as follows.

$$(n-2)\Lambda^{(1)-1/2}S^{(1)}\Lambda^{(1)-1/2} = \sum_{l=1}^{n-2} Z_l Z_l'. \tag{14}$$

Conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\{Z_l\}_{l=1}^{n-2} \quad \perp \quad \hat{\nu}^{(1)}. \tag{15}$$

Conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$Z_l|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \overset{i.i.d}{\sim} N(0, \Lambda^{(1)-1/2}\Sigma^{(1)}\Lambda^{(1)-1/2}), \quad l = 1, \ldots, n-2. \tag{16}$$

For each $l = 1, \ldots, n-2$, we write the vector $Z_l = (\tilde{Z}_{l1}', \ldots, \tilde{Z}_{lp_n}')' \in \mathbb{R}^{p_n s_n}$ with sub-vectors $\tilde{Z}_{lj} = (Z_{lj1}, \ldots, Z_{ljs_n})' \in \mathbb{R}^{s_n}$. Similarly, for each $l = 1, \ldots, n-2$, we let $Z_{l,T} = (\tilde{Z}_{l1}', \ldots, \tilde{Z}_{lq_n}')' \in \mathbb{R}^{q_n s_n}$ and $Z_{l,N} = (\tilde{Z}_{l,q_n+1}', \ldots, \tilde{Z}_{lp_n}')' \in \mathbb{R}^{(p_n-q_n)s_n}$. Accordingly, we denote

$$\begin{aligned}Z_T =&[Z_{1,T}, \ldots, Z_{n-2,T}] \in \mathbb{R}^{q_n s_n \times (n-2)},\\ Z_N =&[Z_{1,N}, \ldots, Z_{n-2,N}] \in \mathbb{R}^{(p_n-q_n)s_n \times (n-2)},\\ Z =&[Z_T', Z_N']' = [Z_1, \ldots, Z_{n-2}] \in \mathbb{R}^{p_n s_n \times (n-2)}.\end{aligned} \tag{17}$$

It follows from (15) and (17) that conditional on nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$Z \quad \perp \quad \hat{\nu}^{(1)}. \tag{18}$$

Based on (14) and (17), it can be observed that

$$\begin{aligned}(n-2)\Lambda_N^{(1)-1/2} S_{NT}^{(1)} \Lambda_T^{(1)-1/2} = Z_N Z_T' = \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T Z_T' \\ + (Z_N - \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T) Z_T'.\end{aligned} \tag{19}$$

The terms $Z_N - \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T$ and $Z_T$ can be expressed as

$$\begin{aligned} Z_N - \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T =& [W Z_1, \ldots, W Z_{n-2}], \\ Z_T =& [W^* Z_1, \ldots, W^* Z_{n-2}], \end{aligned} \tag{20}$$

where

$$\begin{aligned} W &= [-\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2}, I_{(p_n - q_n) s_n}] \in \mathbb{R}^{(p_n - q_n) s_n \times p_n s_n}, \\ W^* &= [I_{q_n s_n}, 0_{q_n s_n \times (p_n - q_n) s_n}] \in \mathbb{R}^{q_n s_n \times p_n s_n}. \end{aligned}$$

Based on (16) and (20), it can be deduced that

$$\begin{bmatrix} W Z_l \\ W^* Z_l \end{bmatrix} \Big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \overset{i.i.d}{\sim} \tag{21}$$

$$N\left(0_{p_n s_n \times 1}, \begin{bmatrix} \Lambda_N^{(1)-1/2} (\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TN}^{(1)}) \Lambda_N^{(1)-1/2} & 0_{(p_n - q_n) s_n \times q_n s_n} \\ 0_{q_n s_n \times (p_n - q_n) s_n} & \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} \end{bmatrix}\right),$$

for $l = 1, \ldots, n-2$. Hence, by combining (16), (20) with (21), it can be concluded that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$Z_T \quad \perp \quad Z_N - \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T. \tag{22}$$

Note that (18) entails that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\begin{aligned} \hat{\nu}_T^{(1)} &\perp \{Z_T, Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T\}, \\ \hat{\nu}_T^{(1)} &\perp Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T, \\ \hat{\nu}_T^{(1)} &\perp Z_T. \end{aligned} \tag{23}$$

Piecing (22) and (23) together yields that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\{\hat{\nu}_T^{(1)}, Z_T\} \perp Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T. \tag{24}$$

In a similar fashion, the quantity $\Lambda_N^{(1)-1/2}\hat{\nu}_N^{(1)}$ can be decomposed into

$$\Lambda_N^{(1)-1/2}\hat{\nu}_N^{(1)} = \Lambda_N^{(1)-1/2}(\hat{\nu}_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}) + \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}. \tag{25}$$

It is not difficult to verify that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\begin{bmatrix} \Lambda_N^{(1)-1/2}(\hat{\nu}_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}) \\ \\ \Lambda_T^{(1)-1/2}\hat{\nu}_T^{(1)} \end{bmatrix} \Bigg| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \sim N\Bigg( \begin{bmatrix} \Lambda_N^{(1)-1/2}(\nu_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}) \\ \\ \Lambda_T^{(1)-1/2}\nu_T^{(1)} \end{bmatrix}, \tag{26}$$

$$nn_1^{-1}n_2^{-1} \begin{bmatrix} \Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2} & 0_{(p_n-q_n)s_n\times q_ns_n} \\ \\ 0_{q_ns_n\times(p_n-q_n)s_n} & \Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2} \end{bmatrix}\Bigg),$$

which further entails that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\hat{\nu}_T^{(1)} \perp \hat{\nu}_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}. \tag{27}$$

Based on (18), it is seen that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\begin{aligned} Z_T &\perp \{\hat{\nu}_T^{(1)}, \hat{\nu}_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\}, \\ Z_T &\perp \hat{\nu}_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}, \\ Z_T &\perp \hat{\nu}_T^{(1)}. \end{aligned} \tag{28}$$

Together with (27) yields that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$\{\hat{\nu}_T^{(1)}, Z_T\} \perp \hat{\nu}_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}. \tag{29}$$

Moreover, using (19) and (25), elementary algebra yields that

$$\Psi = \Pi_1 - \Pi_2 - \Pi_3 - \Pi_4, \tag{30}$$

with

$$\begin{aligned} \Pi_1 &= (n-2)^{-1}(Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T)Z_T'\Lambda_T^{(1)1/2}\tilde{v}_T, \\ \Pi_2 &= \hat{\vartheta}\Lambda_N^{(1)-1/2}(\hat{\nu}_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}), \\ \Pi_3 &= \lambda_n\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}(\Lambda_T^{(1)-1/2}\hat{\Lambda}_T^{(1)1/2} - I_{q_n s_n})\mathrm{sgn}(\beta_T^{(1)}), \\ \Pi_4 &= \lambda_n\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}), \end{aligned}$$

where

$$\begin{aligned} \hat{\vartheta} =& \{n_1 n_2 n^{-1}(n-2)^{-1}\}\{1 + \lambda_n \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\} \\ & \left[1 + \{n_1 n_2 n^{-1}(n-2)^{-1}\}\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\right]^{-1}. \end{aligned}$$

Similar arguments as in the proof of Lemma 11 indicates that there exist universal constants $c_7 > 0$ and $c_9 > c_8 > 0$ such that with probability at least $1 - c_7[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$c_8\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1} \le \hat{\vartheta} \le c_9\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1}. \tag{31}$$

For the term $\Pi_1$, it can be decomposed into

$$\Pi_1 = \Upsilon_1 - \Upsilon_2, \tag{32}$$

where

$$\Upsilon_1 = \hat{\vartheta}(n-2)^{-1}(Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T)Z_T'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)},$$

$$\Upsilon_2 = \lambda_n(n-2)^{-1}(Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T)Z_T'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}).$$

At this point, we denote $\{e_j\}_{j=1}^{(p_n-q_n)s_n}$ as the standard basis in $\mathbb{R}^{(p_n-q_n)s_n}$. Moreover, according to (20), (21) and (24), it can be deduced that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\}$ and for any $j \le (p_n - q_n)s_n$,

$$(Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T)'e_j \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\}$$
$$\sim N\big(0_{(n-2)\times 1}, \{e_j'\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}e_j\}I_{n-2}\big),$$

which implies that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\}$ and for any $j \le (p_n - q_n)s_n$,

$$e_j'\Upsilon_1 \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\} \sim N(0, \Gamma_j),$$

with each

$$\Gamma_j = \hat{\vartheta}^2(n-2)^{-1}\{e_j'\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}e_j\}\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}$$
$$\le \hat{\vartheta}^2(n-2)^{-1}\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}.$$

Together with the maximal inequality, we have that for any $t \ge 0$,

$$P\big[\|\Upsilon_1\|_\infty \ge t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\}\big]$$
$$\le 2(p_n - q_n)s_n \exp\big[-4^{-1}\hat{\vartheta}^{-2}\{\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\}^{-1}nt^2\big].$$

Plugging $t = [8\hat{\vartheta}^2 \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} \log\{(p_n - q_n)s_n\}/n]^{1/2}$ into the above inequality yields

$$P\Big(\|\Upsilon_1\|_\infty \le [8\hat{\vartheta}^2 \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} \log\{(p_n - q_n)s_n\}/n]^{1/2} \Big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\}\Big) \ge 1 - 2\{(p_n - q_n)s_n\}^{-1}. \tag{33}$$

Accordingly, we have

$$\begin{aligned} & P\big(\|\Upsilon_1\|_\infty \le [8\hat{\vartheta}^2 \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} \log\{(p_n - q_n)s_n\}/n]^{1/2}\big) \qquad (34) \\ \ge & \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} \Big\{ \int_{\hat{\nu}_T^{(1)}} \int_{Z_T} f(\hat{\nu}_T^{(1)}, Z_T | \{Y_i = y_i\}_{i=1}^n) \cdot P\big(\|\Upsilon_1\|_\infty \le \\ & [8\hat{\vartheta}^2 \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} \log\{(p_n - q_n)s_n\}/n]^{1/2} \big| \{Y_i = y_i\}_{i=1}^n \cap \{\hat{\nu}_T^{(1)}, Z_T\}\big) \\ & d\hat{\nu}_T^{(1)} dZ_T \Big\} \cdot P\big[\{Y_i = y_i\}_{i=1}^n\big] \ge [1 - 2\{(p_n - q_n)s_n\}^{-1}] \cdot P(\mathcal{M}_n) \\ \ge & 1 - c_{10}[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)], \end{aligned}$$

for some universal constant $c_{10} > 0$, where $f(\hat{\nu}_T^{(1)}, Z_T | \{Y_i = y_i\}_{i=1}^n)$ denotes the conditional joint density function, the second inequality follows from (33), and the last inequality holds from Lemma 3. Based on (31), (34) and Lemma 4, it is seen that there exist universal constants $c_{11}, c_{12} > 0$ that with probability at least $1 - c_{11}[\{(p_n - q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\|\Upsilon_1\|_\infty \le c_{12}[(\sum_{j \in T} \sum_{k=1}^{s_n} \omega_{jk} \beta_{jk}^2)^{-1} \log\{(p_n - q_n)s_n\}/n]^{1/2} \le 2c_{12} K_1^{-1/2} \lambda_n,$$

where the last inequality is by condition (C5). By choosing $K_1 \ge 1600 c_{12}^2 \gamma^{-2}$ in condition (C5), it follows from (C5) and the above inequality that with probability at least $1 - c_{11}[\{(p_n - q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\|\Upsilon_1\|_\infty \le 20^{-1} \gamma \lambda_n. \tag{35}$$

For the term $\Upsilon_2$, similar argument leads to the fact that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\}$ and for any $j \le (p_n - q_n)s_n$,

$$e_j' \Upsilon_2 \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\} \sim N(0, \Xi_j),$$

with each

$$\begin{aligned}\Xi_j =&\lambda_n^2(n-2)^{-1}\{e_j'\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)}-\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}e_j\}\{\mathrm{sgn}(\beta_T^{(1)})' \\ &\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\} \\ \le&\lambda_n^2(n-2)^{-1}\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}.\end{aligned}$$

Together with maximal inequality, we have that for any $t \ge 0$,

$$\begin{aligned}&P\big[\|\Upsilon_2\|_\infty \ge t\big|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\cap\{\hat{\nu}_T^{(1)},Z_T\}\big] \\ \le&2(p_n-q_n)s_n\exp\big[-4^{-1}\lambda_n^{-2}\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1}nt^2\big].\end{aligned}$$

Setting $t=[8\lambda_n^2\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\log\{(p_n-q_n)s_n\}/n]^{1/2}$ in the above inequality yields

$$\begin{aligned}&P\Big(\|\Upsilon_2\|_\infty \le [8\lambda_n^2\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}]^{1/2} \\ &\quad[\log\{(p_n-q_n)s_n\}/n]^{1/2}\Big|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\cap\{\hat{\nu}_T^{(1)},Z_T\}\Big) \\ &\quad\ge 1-2\{(p_n-q_n)s_n\}^{-1}.\end{aligned}$$

Together with similar reasoning as in (34), one has

$$\begin{aligned}&P\big(\|\Upsilon_2\|_\infty \le [8\lambda_n^2\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}]^{1/2} \\ &\quad[\log\{(p_n-q_n)s_n\}/n]^{1/2}\big) \\ &\ge 1-c_{13}[\{(p_n-q_n)s_n\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)],\end{aligned}$$

for some universal constant $c_{13} > 0$. Then, it follows from the above inequality and Lemma 9 that there exist universal constants $c_{14}, c_{15} > 0$ such that with probability at least $1-c_{14}[\{(p_n-q_n)s_n\}^{-1}+(q_ns_n)^{-1}+\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)]$,

$$\|\Upsilon_2\|_\infty \le c_{15}[q_ns_n\log\{(p_n-q_n)s_n\}/n]^{1/2}\lambda_n.$$

Together with (35) and (32), it can be seen that there exist universal constants $c_{16} > 0$ and $c_{17} > 0$ such that with probability at least $1 - c_{16}[\{(p_n - q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\|\Pi_1\|_\infty \leq c_{17}[q_n s_n \log\{(p_n - q_n)s_n\}/n]^{1/2}\lambda_n + 20^{-1}\gamma\lambda_n. \tag{36}$$

For the term $\Pi_2$, (26) together with (29) indicates that

$$\Pi_2\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\} \sim$$
$$N\big(0_{(p_n-q_n)s_n\times 1}, nn_1^{-1}n_2^{-1}\hat{\vartheta}^2\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}\big).$$

Together with the maximal inequality, it can be deduced that for any $t \geq 0$,

$$P(\|\Pi_2\|_\infty \geq t|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\})$$
$$\leq 2(p_n - q_n)s_n \exp\{-(9\pi_1\pi_2\hat{\vartheta}^2)^{-1}nt^2\}.$$

Plugging $t = [18\pi_1\pi_2\hat{\vartheta}^2 \log\{(p_n - q_n)s_n\}/n]^{1/2}$ into the above inequality yields

$$P(\|\Pi_2\|_\infty \leq [18\pi_1\pi_2\hat{\vartheta}^2 \log\{(p_n - q_n)s_n\}/n]^{1/2}$$
$$|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T^{(1)}, Z_T\}) \geq 1 - 2\{(p_n - q_n)s_n\}^{-1}.$$

Together with similar reasoning as in (34), one can show that

$$P(\|\Pi_2\|_\infty \leq [18\pi_1\pi_2\hat{\vartheta}^2 \log\{(p_n - q_n)s_n\}/n]^{1/2})$$
$$\geq 1 - c_{18}[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)],$$

for some constant $c_{18} > 0$. Together with (31), there exist constants $c_{19}, c_{20} > 0$ that with probability at least $1 - c_{19}[\{(p_n - q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\|\Pi_2\|_\infty \leq c_{20}\big(\sum_{j\in T}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^2\big)^{-1/2}\lambda_n. \tag{37}$$

For the term $\Pi_3$, it follows from condition (C2) and Lemma 5 that there exist universal constants $c_{21}, c_{22} > 0$ such that with probability at least $1 - c_{21}\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$, we have $\|\Pi_3\|_\infty \le c_{22}\{q_n s_n \log(q_n s_n)/n\}^{1/2}\lambda_n$. Together with (37), (36) and (30), there exists a universal constant $c_{23} > 0$ such that with probability at least $1 - c_{23}[\{(p_n - q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$, we have $\|\Psi\|_\infty \le (1 - \gamma/2)\lambda_n$. Together with (12) and Lemma 6, the assertion (11) holds trivially, which completes the proof of property (i). To show property (iii), we recall that $\tilde{v} = (\tilde{v}_T', 0')' \in \mathbb{R}^{p_n s_n}$, where $\tilde{v}_T$ is defined in Lemma 2. Together with (9), we have that there exists a universal constants $c_{24} > 0$ such that with probability at least $1 - c_{24}\{(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$,

$$R^\diamond(\hat{v}) = R^\diamond(\tilde{v}) = \pi_1\Omega_1 + \pi_2\Omega_2, \tag{38}$$

where

$$\begin{aligned}\Omega_1 &= \Phi\Big(\big[-\tilde{v}_T'(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)}) - 2^{-1}\tilde{v}_T'\hat{\nu}_T^{(1)} + \{\tilde{v}_T' S_{TT}^{(1)}\tilde{v}_T\}\{\tilde{v}_T'\hat{\nu}_T^{(1)}\}^{-1}\{\log(n_2/n_1)\}\big] \\ &\qquad \{\tilde{v}_T'\Sigma_{TT}^{(1)}\tilde{v}_T\}^{-1/2}\Big), \\ \Omega_2 &= \Phi\Big(\big[-\tilde{v}_T'(\mu_{2,T}^{(1)} - \hat{\mu}_{2,T}^{(1)}) - 2^{-1}\tilde{v}_T'\hat{\nu}_T^{(1)} - \{\tilde{v}_T' S_{TT}^{(1)}\tilde{v}_T\}\{\tilde{v}_T'\hat{\nu}_T^{(1)}\}^{-1}\{\log(n_2/n_1)\}\big] \\ &\qquad \{\tilde{v}_T'\Sigma_{TT}^{(1)}\tilde{v}_T\}^{-1/2}\Big).\end{aligned}$$

Also recalling from (11) of the main paper that

$$R^\circ(\beta^{(1)}) = \pi_1\Omega_1^* + \pi_2\Omega_2^*, \tag{39}$$

with

$$\begin{aligned}\Omega_1^* &= \Phi\Big(-2^{-1}\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{1/2} + \log(\pi_2/\pi_1)\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1/2}\Big), \\ \Omega_2^* &= \Phi\Big(-2^{-1}\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{1/2} - \log(\pi_2/\pi_1)\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1/2}\Big).\end{aligned}$$

We denote $a_n$, $b_n$, $X_n$ and $U_n$ as

$$a_n = 4^{-1}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}, \qquad b_n = \log(\pi_2/\pi_1)\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1/2},$$
$$X_n = \{2\tilde{v}_T'(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)}) + \tilde{v}_T'\hat{\nu}_T^{(1)}\}\{\tilde{v}_T'\Sigma_{TT}^{(1)}\tilde{v}_T\}^{-1/2}\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1/2} - 1,$$
$$U_n = \log(n_2/n_1)\{\tilde{v}_T'\hat{\nu}_T^{(1)}\}^{-1}\{\tilde{v}_T'S_{TT}^{(1)}\tilde{v}_T\}\{\tilde{v}_T'\Sigma_{TT}^{(1)}\tilde{v}_T\}^{-1/2}.$$

Elementary algebra shows that

$$\Omega_1 = \Phi(-a_n^{1/2}(1 + X_n) + U_n), \qquad \Omega_1^* = \Phi(-a_n^{1/2} + b_n). \tag{40}$$

Moreover, under conditions (C2) and (C5), we have

$$a_n \to \infty, \qquad b_n \to 0. \tag{41}$$

Simple algebra indicates that the term $\tilde{v}_T$ can be expressed as

$$\tilde{v}_T = \hat{\vartheta}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \lambda_n S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}), \tag{42}$$

where $\hat{\vartheta} = \{n_1n_2n^{-1}(n-2)^{-1}\}\{1 + \lambda_n\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})\} \cdot \big[1 + \{n_1n_2n^{-1}(n-2)^{-1}\}\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\big]^{-1}$. We further define $\tilde{\vartheta}$ as

$$\begin{aligned}\tilde{\vartheta} =& \{n_1n_2n^{-1}(n-2)^{-1}\}\{1 + \lambda_n\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})\}\\ &\cdot \big[1 + \{n_1n_2n^{-1}(n-2)^{-1}\}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\big]^{-1}.\end{aligned}$$

It then follows from Lemma 3, Lemma 4, Lemma 10, (93) and (41) that there exist universal constants $c_{25}, c_{26} > 0$ such that with probability at least $1 - c_{25}[(q_ns_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\begin{aligned}&|\hat{\vartheta} - \tilde{\vartheta}| \le c_{26}\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1}[q_ns_n/n + \{\log\log(n)/n\}^{1/2}] + c_{26}\lambda_n\{\nu_T^{(1)'}\cdot\\ &\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1/2}[(q_ns_n)^{3/2}/n + \{q_ns_n\log(q_ns_n)/n\}^{1/2} + \{q_ns_n\log\log(n)/n\}^{1/2}].\end{aligned} \tag{43}$$

For the term $\tilde{v}_T' S_{TT}^{(1)} \tilde{v}_T$, using (42), we have

$$\tilde{v}_T' S_{TT}^{(1)} \tilde{v}_T = \tilde{\vartheta}^2 \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} + \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3, \tag{44}$$

where $\mathcal{I}_1 = \hat{\vartheta}^2 \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \tilde{\vartheta}^2 \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}$, $\mathcal{I}_2 = \lambda_n^2 \mathrm{sgn}(\beta_T^{(1)})' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})$, $\mathcal{I}_3 = -2\hat{\vartheta} \lambda_n \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})$. For the term $\mathcal{I}_1$, since $|\mathcal{I}_1| \leq \hat{\vartheta}^2 |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| + |\hat{\vartheta} - \tilde{\vartheta}| \cdot (2|\hat{\vartheta}| + |\hat{\vartheta} - \tilde{\vartheta}|) \cdot \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}$, it follows from Lemma 4, (31), (41), and (43) that there exist constants $c_{27}, c_{28} > 0$ such that with probability at least $1 - c_{27}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\begin{aligned} |\mathcal{I}_1| \leq & c_{28} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1} [q_n s_n / n + \{\log\log(n)/n\}^{1/2}] + c_{28} \lambda_n \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1/2} \cdot \\ & [(q_n s_n)^{3/2}/n + \{q_n s_n \log(q_n s_n)/n\}^{1/2} + \{q_n s_n \log\log(n)/n\}^{1/2}]. \end{aligned}$$

To bound the term $\mathcal{I}_2$, since $|\mathcal{I}_2| \lesssim \lambda_n^2 q_n s_n \big[1 + \big|\{\mathrm{sgn}(\beta_T^{(1)})' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})\} \cdot \{\mathrm{sgn}(\beta_T^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})\}^{-1} - 1\big|\big]$, it follows from Lemma 9 that there exist universal constants $c_{29}, c_{30} > 0$ such that with probability at least $1 - c_{29}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\mathcal{I}_2| \leq c_{30} \lambda_n^2 q_n s_n.$$

For the term $\mathcal{I}_3$, since $|\mathcal{I}_3| \leq 2\lambda_n |\hat{\vartheta}| \cdot |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)}) - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| + 2|\hat{\vartheta}| \cdot \{\lambda_n \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})\}$, it follows from Lemma 10, (93) and (31) that there exist constants $c_{31}, c_{32} > 0$ such that with probability at least $1 - c_{31}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\mathcal{I}_3| \leq c_{32} (\lambda_n^2 q_n s_n)^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1/2}.$$

By combining the above three inequalities with (44), we have that there exist universal constants $c_{33}, c_{34} > 0$ such that with probability at least $1 - c_{33}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} +$

$\exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\tilde{\upsilon}_T' S_{TT}^{(1)} \tilde{\upsilon}_T - \tilde{\vartheta}^2 \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| \leq c_{34} (\lambda_n^2 q_n s_n)^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1/2}. \tag{45}$$

Since $|\tilde{\upsilon}_T' S_{TT}^{(1)} \tilde{\upsilon}_T - \tilde{\upsilon}_T' \Sigma_{TT}^{(1)} \tilde{\upsilon}_T| \leq \lambda_{\max}(\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}) \|\Lambda_T^{(1)-1/2} (\Sigma_{TT}^{(1)} - S_{TT}^{(1)}) \Lambda_T^{(1)-1/2}\|_2 \tilde{\upsilon}_T' S_{TT}^{(1)} \tilde{\upsilon}_T$, it follows from Lemma 7 and Lemma 8 that there exist constants $c_{35}, c_{36} > 0$ such that with probability at least $1 - c_{35}\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$,

$$|\tilde{\upsilon}_T' S_{TT}^{(1)} \tilde{\upsilon}_T - \tilde{\upsilon}_T' \Sigma_{TT}^{(1)} \tilde{\upsilon}_T| \leq c_{36} \{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2} \tilde{\upsilon}_T' S_{TT}^{(1)} \tilde{\upsilon}_T. \tag{46}$$

For the term $\tilde{\upsilon}_T' \hat{\nu}_T^{(1)}$, using (42) again, it has the form

$$\tilde{\upsilon}_T' \hat{\nu}_T^{(1)} = \tilde{\vartheta} \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} + \mathcal{V}_1 + \mathcal{V}_2, \tag{47}$$

where $\mathcal{V}_1 = \hat{\vartheta} \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \tilde{\vartheta} \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}$ and $\mathcal{V}_2 = -\lambda_n \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})$. S-ince $|\mathcal{V}_1| \leq |\hat{\vartheta}| \cdot |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| + |\hat{\vartheta} - \tilde{\vartheta}| \cdot \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}$, it follows from Lemma 4, (31), (43) and (41) that there exist universal constants $c_{37}, c_{38} > 0$ such that with probability at least $1 - c_{37}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\begin{aligned}|\mathcal{V}_1| \leq & c_{38}[q_n s_n/n + \{\log\log(n)/n\}^{1/2}] + c_{38} \lambda_n \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2} \\ & \cdot [(q_n s_n)^{3/2}/n + \{q_n s_n \log(q_n s_n)/n\}^{1/2} + \{q_n s_n \log\log(n)/n\}^{1/2}].\end{aligned}$$

Since $|\mathcal{V}_2| \leq \lambda_n |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)}) - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| + \lambda_n \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})$, it holds from Lemma 10, (93), and (41) that there exist constants $c_{39}, c_{40} > 0$ such that with probability at least $1 - c_{39}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\mathcal{V}_2| \leq c_{40} (\lambda_n^2 q_n s_n)^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2}.$$

By combining the above two inequalities with (47), we conclude that there exist universal constants $c_{41}, c_{42} > 0$ such that with probability at least $1 - c_{41}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\tilde{\upsilon}_T' \hat{\nu}_T^{(1)} - \tilde{\vartheta} \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| \leq c_{42} (\lambda_n^2 q_n s_n)^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2}. \tag{48}$$

Moreover, using (31), (45), (46), (48), and the fact that $\lambda_n^2 q_n s_n \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} = o(1)$, elementary calculation indicates that

$$2a_n^{1/2}(U_n - b_n) = o_p(1). \tag{49}$$

For the term $\tilde{v}_T'(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})$, it follows from (42) and Holder's inequality that

$$|\tilde{v}_T'(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})| \le \|\Lambda_T^{(1)-1/2}(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})\|_\infty \{q_n s_n \lambda_{\max}(\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2})\}^{1/2}\cdot$$
$$\big[|\hat{\vartheta}| \cdot \{\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)}\}^{1/2} + \lambda_n \{\mathrm{sgn}(\beta_T^{(1)})' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})\}^{1/2}\big].$$

Together with Lemma 4, Lemma 8, Lemma 9 and (31), it can be deduced that there exist universal constants $c_{43}, c_{44} > 0$ such that with probability at least $1 - c_{43}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\tilde{v}_T'(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})| \le c_{44}(q_n s_n)^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1/2} \|\Lambda_T^{(1)-1/2}(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})\|_\infty. \tag{50}$$

To bound the term $\|\Lambda_T^{(1)-1/2}(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})\|_\infty$, note that

$$\Lambda_T^{(1)-1/2}(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \sim N(0, n_1^{-1} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}).$$

Union bound inequality and the concentration inequality imply that for any $t \ge 0$,

$$P\big\{\|\Lambda_T^{(1)-1/2}(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})\|_\infty \ge t\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big\} \le 2q_n s_n \exp\{-(\pi_1/4)nt^2\}.$$

Plugging $t = c_{45}\{\log(q_n s_n)/n\}^{1/2}$ with $c_{45} = (8/\pi_1)^{1/2}$ into the above yields $P\big[\|\Lambda_T^{(1)-1/2}(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})\|_\infty \le c_{45}\{\log(q_n s_n)/n\}^{1/2}\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \ge 1 - 2(q_n s_n)^{-1}$. Together with Lemma 3, it can be deduced that $P\big[\|\Lambda_T^{(1)-1/2}(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})\|_\infty \le c_{45}\{\log(q_n s_n)/n\}^{1/2}\big] \ge 1 - 2\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$. Together with (50), there exist universal constants $c_{46}, c_{47} > 0$ such that with probability at least $1 - c_{46}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\tilde{v}_T'(\hat{\mu}_{1,T}^{(1)} - \mu_{1,T}^{(1)})| \le c_{47}\{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1/2} \{q_n s_n \log(q_n s_n)/n\}^{1/2}.$$

Together with (31), (45), (46), (48), and conditions (C2)–(C5), it is seen that $4a_nX_n = o_p(1)$. Together with (49), (41), (40), and Lemma 12, it can be concluded that

$$\Omega_1/\Omega_1^* \xrightarrow{p} 1, \qquad \Omega_1^* \to 0. \tag{51}$$

Similar argument leads to $\Omega_2/\Omega_2^* \xrightarrow{p} 1$, $\Omega_2^* \to 0$. Together with (38), (39), and (51), it holds that $R^\diamond(\hat{v})/R^\circ(\beta^{(1)}) \xrightarrow{p} 1$, $R^\circ(\beta^{(1)}) \to 0$, which completes the proof. □

*Proof of Corollary 1*: It follows directly from Theorems 1 and 2. □

In the next section, we present all the auxiliary lemmas with their proofs.

# 3 Auxiliary lemmas with their proofs

**Lemma 1.** *Assume the following conditions (a)–(b):*

*(a)* $c_1 \le \lambda_{\min}(\Lambda^{\dagger 1/2}\Sigma\Lambda^{\dagger 1/2}) \le \lambda_{\max}(\Lambda^{\dagger 1/2}\Sigma\Lambda^{\dagger 1/2}) \le c_2$,

$c_1 \le \lambda_{\min}(\Lambda^{(1)-1/2}\Sigma^{(1)}\Lambda^{(1)-1/2}) \le \lambda_{\max}(\Lambda^{(1)-1/2}\Sigma^{(1)}\Lambda^{(1)-1/2}) \le c_2$,

*for some universal constants* $0 < c_1 < c_2$.

*(b)* $\sum_{j\in T^*}\sum_{k=s_n+1}^{\infty}\omega_{jk}\beta_{jk}^{*2} = o(\min_{j\in T^*}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^{*2})$.

*Then we have the following properties:*

*1)* $N \subseteq N^*$ *and* $T^* \subseteq T$.

*2)* $\Delta^{(1)2} = \{1 + o(r_n^{-1}) + o(r_n^{-1/2}\alpha_n^{1/2})\}\Delta^2$,

*where the parameter* $\alpha_n = (\beta_T^{*(1)'}\Sigma_{TT}^{(1,2)}\Sigma_{TT}^{(2)\dagger}\Sigma_{TT}^{(2,1)}\beta_T^{*(1)})/(\beta_T^{*(1)'}\Sigma_{TT}^{(1)}\beta_T^{*(1)}) \le 1$.

*Proof of Lemma 1*: First of all, we note that the equation $\Sigma\beta^* = \nu$ is equivalent to

$$\begin{bmatrix} \Sigma_{TT}^{(1)} & \Sigma_{TN}^{(1)} & \Sigma_{TT}^{(1,2)} & \Sigma_{TN}^{(1,2)} \\ \Sigma_{NT}^{(1)} & \Sigma_{NN}^{(1)} & \Sigma_{NT}^{(1,2)} & \Sigma_{NN}^{(1,2)} \\ \Sigma_{TT}^{(2,1)} & \Sigma_{TN}^{(2,1)} & \Sigma_{TT}^{(2)} & \Sigma_{TN}^{(2)} \\ \Sigma_{NT}^{(2,1)} & \Sigma_{NN}^{(2,1)} & \Sigma_{NT}^{(2)} & \Sigma_{NN}^{(2)} \end{bmatrix} \cdot \begin{bmatrix} \beta_T^{*(1)} \\ \beta_N^{*(1)} \\ \beta_T^{*(2)} \\ \beta_N^{*(2)} \end{bmatrix} = \begin{bmatrix} \nu_T^{(1)} \\ \nu_N^{(1)} \\ \nu_T^{(2)} \\ \nu_N^{(2)} \end{bmatrix}$$

which entails that

$$\Sigma_{TT}^{(1)}\beta_T^{*(1)} + \Sigma_{TN}^{(1)}\beta_N^{*(1)} + \Sigma_{TT}^{(1,2)}\beta_T^{*(2)} + \Sigma_{TN}^{(1,2)}\beta_N^{*(2)} = \nu_T^{(1)}, \tag{52}$$

$$\Sigma_{NT}^{(1)}\beta_T^{*(1)} + \Sigma_{NN}^{(1)}\beta_N^{*(1)} + \Sigma_{NT}^{(1,2)}\beta_T^{*(2)} + \Sigma_{NN}^{(1,2)}\beta_N^{*(2)} = \nu_N^{(1)}. \tag{53}$$

Multiplying both sides of (52) by $\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}$, we obtain

$$\Sigma_{NT}^{(1)}\beta_T^{*(1)} + \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}\beta_N^{*(1)} + \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)} + \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)}\beta_N^{*(2)} = \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}.$$

By subtracting (53) from the above equation, it can be seen that

$$\begin{aligned} &(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)} + (\Sigma_{NT}^{(1,2)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)})\beta_T^{*(2)} + \\ &(\Sigma_{NN}^{(1,2)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)})\beta_N^{*(2)} = \nu_N^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}. \end{aligned}$$

By combining (10) in the main paper with the above equation, it can be deduced that

$$\begin{aligned} &\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)} \\ =&\Lambda_N^{(1)-1/2}(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)} - \Sigma_{NT}^{(1,2)})\beta_T^{*(2)} + \Lambda_N^{(1)-1/2}(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)} - \Sigma_{NN}^{(1,2)})\beta_N^{*(2)}. \end{aligned}$$

Together with the triangle inequality, we have

$$\begin{aligned}
&\|\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)}\|_2 \\
\leq&\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}\|_2 + \|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)}\beta_T^{*(2)}\|_2 + \\
&\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)}\beta_N^{*(2)}\|_2 + \|\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}\beta_N^{*(2)}\|_2,
\end{aligned}$$

which further implies that

$$\begin{aligned}
&\|\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)}\|_2^2 \\
\lesssim&\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}\|_2^2 + \|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)}\beta_T^{*(2)}\|_2^2 + \\
&\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)}\beta_N^{*(2)}\|_2^2 + \|\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}\beta_N^{*(2)}\|_2^2. \qquad (54)
\end{aligned}$$

Based on condition (a) and Lemma 14, it is trivial to show that

$$\begin{aligned}
&\lambda_{\min}\big(\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}\big) \\
=&\lambda_{\max}^{-1}\big(\Lambda_N^{(1)1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})^{-1}\Lambda_N^{(1)1/2}\big) \geq c_1. \qquad (55)
\end{aligned}$$

for the universal constant $c_1 > 0$ defined in condition (a). Hence, for the term $\|\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)}\|_2^2$, we have

$$\begin{aligned}
&\|\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)}\|_2^2 \\
\geq&c_1\lambda_{\max}\big(\Lambda_N^{(1)1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})^{-1}\Lambda_N^{(1)1/2}\big) \cdot \{(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)}\}' \\
&\cdot \Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2}\{(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\beta_N^{*(1)}\} \\
\geq&c_1(\Lambda_N^{(1)1/2}\beta_N^{*(1)})'\{\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}\}(\Lambda_N^{(1)1/2}\beta_N^{*(1)}) \\
\geq&c_1^2\|\Lambda_N^{(1)1/2}\beta_N^{*(1)}\|_2^2, \qquad (56)
\end{aligned}$$

where the first inequality is by (55), and the last inequality is also based on (55). According to condition (a) and Lemma 14 again, we have

$$\lambda_{\min}\big(\Lambda_N^{(1)1/2}\Sigma_{NN}^{(1)-1}\Lambda_N^{(1)1/2}\big) = \lambda_{\max}^{-1}\big(\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1)}\Lambda_N^{(1)-1/2}\big) \geq c_2^{-1}, \tag{57}$$

$$\lambda_{\min}\big(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}\big) \geq c_1, \tag{58}$$

$$\lambda_{\max}\big(\Lambda_T^{(1)-1/2}\Sigma_{TN}^{(1)}\Sigma_{NN}^{(1)-1}\Sigma_{NT}^{(1)}\Lambda_T^{(1)-1/2}\big) \leq \lambda_{\max}\big(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}\big) \leq c_2, \tag{59}$$

$$\lambda_{\max}\big(\Lambda_T^{(2)\dagger1/2}\Sigma_{TT}^{(2,1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\Lambda_T^{(2)\dagger1/2}\big) \leq \lambda_{\max}\big(\Lambda_T^{(2)\dagger1/2}\Sigma_{TT}^{(2)}\Lambda_T^{(2)\dagger1/2}\big) \leq c_2, \tag{60}$$

for the universal constants $c_1$ and $c_2$ defined in condition (a). Thus, for the term $\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}\|_2^2$, we have

$$\begin{aligned}
&\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}\|_2^2 \\
\leq& c_2\lambda_{\min}\big(\Lambda_N^{(1)1/2}\Sigma_{NN}^{(1)-1}\Lambda_N^{(1)1/2}\big)(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)})'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}) \\
\leq& c_2(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)})'\Sigma_{NN}^{(1)-1}(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}) \\
\leq& c_2\lambda_{\max}\big(\Lambda_T^{(1)-1/2}\Sigma_{TN}^{(1)}\Sigma_{NN}^{(1)-1}\Sigma_{NT}^{(1)}\Lambda_T^{(1)-1/2}\big)\|\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}\|_2^2 \\
\leq& c_2^2(\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)})'(\Lambda_T^{(1)1/2}\Lambda_T^{(1)1/2})(\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}) \\
\leq& c_2^2c_1^{-1}\lambda_{\min}\big(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}\big)(\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)})'(\Lambda_T^{(1)1/2}\Lambda_T^{(1)1/2})(\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}) \\
\leq& c_2^2c_1^{-1}\lambda_{\max}\big(\Lambda_T^{(2)\dagger1/2}\Sigma_{TT}^{(2,1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\Lambda_T^{(2)\dagger1/2}\big)\|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2^2 \\
\leq& c_2^3c_1^{-1}\|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2^2,
\end{aligned}$$

where the first inequality is by (57), the fourth inequality follows from (59), the fifth inequality is based on (58), and the last inequality is according to (60). Likewise, for the term $\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)}\beta_T^{*(2)}\|_2^2$, we have

$$\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)}\beta_T^{*(2)}\|_2^2 \leq c_2^2\|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2^2.$$

In a similar fashion, for the term $\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)}\beta_N^{*(2)}\|_2^2$, we have

$$\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)}\beta_N^{*(2)}\|_2^2 \leq c_2^3c_1^{-1}\|\Lambda_N^{(2)1/2}\beta_N^{*(2)}\|_2^2.$$

In addition, for the term $\|\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}\beta_N^{*(2)}\|_2^2$, one has

$$
\begin{aligned}
&\|\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}\beta_N^{*(2)}\|_2^2 = (\Sigma_{NN}^{(1,2)}\beta_N^{*(2)})'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NN}^{(1,2)}\beta_N^{*(2)}) \\
\leq& c_2\lambda_{\min}\big(\Lambda_N^{(1)1/2}\Sigma_{NN}^{(1)-1}\Lambda_N^{(1)1/2}\big)(\Sigma_{NN}^{(1,2)}\beta_N^{*(2)})'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NN}^{(1,2)}\beta_N^{*(2)}) \\
\leq& c_2(\Lambda_N^{(2)1/2}\beta_N^{*(2)})'(\Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2,1)}\Sigma_{NN}^{(1)-1}\Sigma_{NN}^{(1,2)}\Lambda_N^{(2)\dagger 1/2})(\Lambda_N^{(2)1/2}\beta_N^{*(2)}) \\
\leq& c_2\lambda_{\max}\big(\Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2,1)}\Sigma_{NN}^{(1)-1}\Sigma_{NN}^{(1,2)}\Lambda_N^{(2)\dagger 1/2}\big)\|\Lambda_N^{(2)1/2}\beta_N^{*(2)}\|_2^2 \\
\leq& c_2\lambda_{\max}\big(\Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2)}\Lambda_N^{(2)\dagger 1/2}\big)\|\Lambda_N^{(2)1/2}\beta_N^{*(2)}\|_2^2 \\
\leq& c_2^2\|\Lambda_N^{(2)1/2}\beta_N^{*(2)}\|_2^2.
\end{aligned}
$$

To this end, based on the above four inequalities, (56) and (54), we conclude that

$$
\|\Lambda_N^{(1)1/2}\beta_N^{*(1)}\|_2^2 \lesssim \|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2^2 + \|\Lambda_N^{(2)1/2}\beta_N^{*(2)}\|_2^2,
$$

which is equivalent to

$$
\sum_{j\in N}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^{*2} \lesssim \sum_{j\in T^*}\sum_{k=s_n+1}^{\infty}\omega_{jk}\beta_{jk}^{*2}.
$$

Together with condition (b), it can be derived that

$$
\sum_{j\in N}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^{*2} = o(\min_{j\in T^*}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^{*2}),
$$

entailing $N \subseteq N^*$ and $T^* \subseteq T$, which completes the proof of 1). To prove property 2), by substituting $\beta_N^* = 0$ into (52), we obtain the equation

$$
\beta_T^{(1)} - \beta_T^{*(1)} = \Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}. \tag{61}
$$

Moreover, we note that

$$
\begin{aligned}
\Delta^2 &= \nu_{T^*}'\Sigma_{T^*T^*}^{\dagger}\nu_{T^*} = \beta_{T^*}^{*'}\Sigma_{T^*T^*}\beta_{T^*}^* = \beta_T^{*'}\Sigma_{TT}\beta_T^* = \nu_T'\Sigma_{TT}^{\dagger}\nu_T \\
&= \beta_T^{*(1)'}\Sigma_{TT}^{(1)}\beta_T^{*(1)} + 2\beta_T^{*(1)'}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)} + \beta_T^{*(2)'}\Sigma_{TT}^{(2)}\beta_T^{*(2)},
\end{aligned} \tag{62}
$$

where the third equality follows from $T^* \subseteq T$. For the term $\beta_T^{*(2)'}\Sigma_{TT}^{(2)}\beta_T^{*(2)}$, we have

$$\begin{aligned}
&\beta_T^{*(2)'}\Sigma_{TT}^{(2)}\beta_T^{*(2)} \leq \lambda_{\max}\big(\Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2)}\Lambda_T^{(2)\dagger 1/2}\big)\|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2^2 \\
\leq & c_2\sum_{j\in T^*}\sum_{k=s_n+1}^{\infty}\omega_{jk}\beta_{jk}^{*2} \leq c_2 o(\min_{j\in T^*}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^{*2}) \leq c_2 r_n^{-1} o(\sum_{j\in T^*}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^{*2}) \\
\leq & c_2 r_n^{-1} o(\sum_{j\in T^*}\sum_{k=1}^{\infty}\omega_{jk}\beta_{jk}^{*2}) \leq \lambda_{\min}\big(\Lambda_{T^*}^{\dagger 1/2}\Sigma_{T^*T^*}\Lambda_{T^*}^{\dagger 1/2}\big)(\beta_{T^*}^{*'}\Lambda_{T^*}^{1/2}\Lambda_{T^*}^{1/2}\beta_{T^*}^{*})o(r_n^{-1}) \\
\leq & (\beta_{T^*}^{*'}\Sigma_{T^*T^*}\beta_{T^*}^{*})o(r_n^{-1}) \leq \Delta^2 o(r_n^{-1}),
\end{aligned} \tag{63}$$

where the last inequality is by (62). Regarding the term $\beta_T^{*(1)'}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}$, one has

$$|\beta_T^{*(1)'}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}| \leq \|\Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)}\beta_T^{*(1)}\|_2\|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2.$$

For the term $\|\Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)}\beta_T^{*(1)}\|_2$, we have

$$\begin{aligned}
&\|\Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)}\beta_T^{*(1)}\|_2^2 \lesssim \beta_T^{*(1)'}(\Sigma_{TT}^{(1,2)}\Sigma_{TT}^{(2)\dagger}\Sigma_{TT}^{(2,1)})\beta_T^{*(1)} \lesssim \alpha_n\beta_T^{*(1)'}\Sigma_{TT}^{(1)}\beta_T^{*(1)} \\
\lesssim & \alpha_n\|\Lambda_{T^*}^{1/2}\beta_{T^*}^{*}\|_2^2 \lesssim \alpha_n\beta_{T^*}^{*'}\Sigma_{T^*T^*}\beta_{T^*}^{*} \lesssim \alpha_n\Delta^2,
\end{aligned}$$

where the last inequality is by (62). For the term $\|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2$, one has

$$\|\Lambda_T^{(2)1/2}\beta_T^{*(2)}\|_2^2 \lesssim \|\Lambda_{T^*}^{(2)1/2}\beta_{T^*}^{*(2)}\|_2^2 \lesssim o(\min_{j\in T^*}\sum_{k=1}^{s_n}\omega_{jk}\beta_{jk}^{*2}) \lesssim \sum_{j\in T^*}\sum_{k=1}^{\infty}\omega_{jk}\beta_{jk}^{*2}o(r_n^{-1}) \lesssim \Delta^2 o(r_n^{-1}).$$

To this end, based on the above three inequalities, we have

$$|\beta_T^{*(1)'}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)}| \lesssim \Delta^2 o(r_n^{-1/2}\alpha_n^{1/2}). \tag{64}$$

For the term $\beta_T^{*(1)'}\Sigma_{TT}^{(1)}\beta_T^{*(1)}$, we have

$$\begin{aligned}
\beta_T^{*(1)'}\Sigma_{TT}^{(1)}\beta_T^{*(1)} =& \beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)} - (\beta_T^{*(1)} - \beta_T^{(1)})'\Sigma_{TT}^{(1)}(\beta_T^{*(1)} - \beta_T^{(1)}) + 2\beta_T^{*(1)'}\Sigma_{TT}^{(1)}(\beta_T^{*(1)} - \beta_T^{(1)}) \\
=& \beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)} - \beta_T^{*(2)'}\Sigma_{TT}^{(2,1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)} - 2\beta_T^{*(1)'}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)} \\
=& \beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)} + O(1)\beta_T^{*(2)'}\Sigma_{TT}^{(2)}\beta_T^{*(2)} - 2\beta_T^{*(1)'}\Sigma_{TT}^{(1,2)}\beta_T^{*(2)} \\
=& \beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)} + \Delta^2 o(r_n^{-1} + r_n^{-1/2}\alpha_n^{1/2}),
\end{aligned}$$

where the second equality follows from (61), and the last equality is based on (63) and (64). Together with (64), (63) and (62), it can be concluded that $\Delta^{(1)2} = \{1 + o(r_n^{-1}) + o(r_n^{-1/2}\alpha_n^{1/2})\}\Delta^2$, which completes the proof. □

**Lemma 2.** *Assume the invertibility of $S_{TT}^{(1)}$ and consider the following optimization problem:*

$$\min_{\upsilon_T \in \mathbb{R}^{q_n s_n}} \Big[\frac{1}{2}\upsilon_T'\{S_{TT}^{(1)} + \frac{n_1 n_2}{n(n-2)}\hat{\nu}_T^{(1)}\hat{\nu}_T^{(1)'}\}\upsilon_T - \frac{n_1 n_2}{n(n-2)}\upsilon_T'\hat{\nu}_T^{(1)} + \lambda_n(\hat{\Lambda}_T^{(1)1/2}\upsilon_T)' sgn(\beta_T^{(1)})\Big],$$

*where $\upsilon_T = (\upsilon_1', \ldots, \upsilon_{q_n}')'$ with sub-vectors $\upsilon_j = (\upsilon_{j1}, \ldots, \upsilon_{js_n})' \in \mathbb{R}^{s_n}$. Let $\tilde{\upsilon}_T$ be the solution of this optimization problem where $\tilde{\upsilon}_T = (\tilde{\upsilon}_1', \ldots, \tilde{\upsilon}_{q_n}')'$ with sub-vectors $\tilde{\upsilon}_j = (\tilde{\upsilon}_{j1}, \ldots, \tilde{\upsilon}_{js_n})' \in \mathbb{R}^{s_n}$, then we have:*

$$\begin{aligned}\tilde{\upsilon}_T =& \{n_1 n_2 n^{-1}(n-2)^{-1}\}\{1 + \lambda_n \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2} sgn(\beta_T^{(1)})\} \\ & \cdot \big[1 + \{n_1 n_2 n^{-1}(n-2)^{-1}\}\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\big]^{-1} S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \lambda_n S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2} sgn(\beta_T^{(1)}).\end{aligned}$$

*Proof of Lemma 2*: The proof is analogous to that of Lemma 16. □

**Lemma 3.** *Define the events $\mathcal{M}_n$ and $\mathcal{M}_n^*$ as*

$$\mathcal{M}_n = \{\pi_1/2 \le n_1/n \le 3\pi_1/2\} \cap \{\pi_2/2 \le n_2/n \le 3\pi_2/2\},$$

$$\mathcal{M}_n^* = \{\pi_1\pi_2/4 \le n_1 n_2/n^2 \le 9\pi_1\pi_2/4\}.$$

*Then we have the following properties:*

*1)* $P(\mathcal{M}_n) \ge 1 - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12)$.

*2)* $P(\mathcal{M}_n^*) \ge 1 - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12)$.

*Proof of Lemma 3*: First of all, note that $n_1 \sim \text{Binomial}(n, \pi_1)$. Invoking the chernoff tail bounds for binomial random variables, we have that for any $\delta \in [0, 1]$,

$$P\{n_1 \ge (1+\delta)n\pi_1\} \le \exp(-n\pi_1\delta^2/3),$$

$$P\{n_1 \le (1-\delta)n\pi_1\} \le \exp(-n\pi_1\delta^2/3).$$

Then, we substitute $\delta = 1/2$ into the above two inequalities to obtain

$$\begin{aligned} &P(n_1/n \geq 3\pi_1/2) \leq \exp(-n\pi_1/12), \\ &P(n_1/n \leq \pi_1/2) \leq \exp(-n\pi_1/12). \end{aligned} \tag{65}$$

Accordingly, we have

$$\begin{aligned} P(\pi_1/2 \leq n_1/n \leq 3\pi_1/2) =& 1 - P(n_1/n > 3\pi_1/2) - P(n_1/n < \pi_1/2) \\ \geq& 1 - 2\exp(-n\pi_1/12), \end{aligned}$$

where the last inequality is by (65). By symmetry, one has

$$P(\pi_2/2 \leq n_2/n \leq 3\pi_2/2) \geq 1 - 2\exp(-n\pi_2/12).$$

To this end, based on the above two inequalities, we can deduce that $P(\mathcal{M}_n) \geq 1 - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12)$, which completes the proof of 1). Property 2) follows from the fact that $\mathcal{M}_n \subseteq \mathcal{M}_n^*$. □

**Lemma 4.** *For any $\varrho \in (e^{-n/100}, 1/100)$, define the event $\mathcal{M}_{3n}(\varrho)$ as*

$$\begin{aligned} \mathcal{M}_{3n}(\varrho) =& \Big\{ \big| \hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \big| \lesssim q_n s_n/n + \log(\varrho^{-1})/n \\ &+ \big[ q_n s_n/n + \{\log(\varrho^{-1})/n\}^{1/2} \big] \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\} + \{\log(\varrho^{-1})/n\}^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2} \Big\}. \end{aligned}$$

*Assume the condition (a):*

*(a)* $q_n s_n = o(n)$.

*Then we have the following property:*

$$P\{\mathcal{M}_{3n}(\varrho)\} \geq 1 - 4\varrho - 4\exp(-n\pi_1/12) - 4\exp(-n\pi_2/12), \quad \forall \varrho \in (e^{-n/100}, 1/100).$$

*Proof of Lemma 4*: First of all, note that

$$
\begin{aligned}
&\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \\
=&(\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)})(\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} / \hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - 1) \\
&+ \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} (\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} / \hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - 1) + (\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}),
\end{aligned}
$$

which implies that

$$
\begin{aligned}
&|\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| \\
\leq&|\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| \cdot |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} / \hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - 1| \\
&+ \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} / \hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - 1| + |\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}|.
\end{aligned}
$$

Together with Lemma 18 and Lemma 19, we conclude that with probability at least $1 - 4\varrho - 4\exp(-n\pi_1/12) - 4\exp(-n\pi_2/12)$,

$$
\begin{aligned}
|\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| \lesssim& q_n s_n/n + \log(\varrho^{-1})/n + \left[q_n s_n/n + \{\log(\varrho^{-1})/n\}^{1/2}\right] \\
&\cdot \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\} + \{\log(\varrho^{-1})/n\}^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2},
\end{aligned}
$$

which completes the proof. □

**Lemma 5.** *Assume the following condition (a):*

*(a)* $\log(q_n s_n) = o(n)$.

*Then there exist universal constants* $c_1 > 0$ *and* $c_2 > 0$ *such that:*

*1)* $P\left[\|\hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \leq c_1 \{\log(q_n s_n)/n\}^{1/2}\right] \geq 1 - c_2\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$.

*2)* $P\left[\|\Lambda_T^{(1)} \hat{\Lambda}_T^{(1)-1} - I_{q_n s_n}\|_{\max} \leq c_1 \{\log(q_n s_n)/n\}^{1/2}\right] \geq 1 - c_2\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$.

*3)* $P\big[\|\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} - I_{q_n s_n}\|_{\max} \leq c_1\{\log(q_n s_n)/n\}^{1/2}\big] \geq 1 - c_2\{(q_n s_n)^{-1}$

$+ \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}.$

*4)* $P\big[\|\Lambda_T^{(1)1/2}\hat{\Lambda}_T^{(1)-1/2} - I_{q_n s_n}\|_{\max} \leq c_1\{\log(q_n s_n)/n\}^{1/2}\big] \geq 1 - c_2\{(q_n s_n)^{-1}$

$+ \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}.$

*Note that $I_{q_n s_n}$ denotes the $q_n s_n \times q_n s_n$ identity matrix.*

*Proof of Lemma 5*: Before showing the Lemma, we prepare some notations. For any sub-exponential random variable $X$, its sub-exponential norm is denoted as $\|X\|_\psi = \sup_{q\geq 1} q^{-1}\{E(|X|^q)\}^{1/q}$. Now, we are in a position to start the proof. First of all, notice that

$$\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} = \max_{j\in T}\max_{k\leq s_n}|\hat{\omega}_{jk}\omega_{jk}^{-1} - 1|. \tag{66}$$

Moreover, by definition, we have that for every $j \in T$ and $k \leq s_n$,

$$\begin{aligned}\hat{\omega}_{jk} =&(n-2)^{-1}\Big[n_1\big\{\sum_{i\in H_1}(\xi_{ijk} - \mu_{1jk})^2/n_1\big\} + n_2\big\{\sum_{i'\in H_2}(\xi_{i'jk} - \mu_{2jk})^2/n_2\big\}\Big]\\ &- (n-2)^{-1}\Big[n_1(\sum_{i_1\in H_1}\xi_{i_1jk}/n_1 - \mu_{1jk})^2 + n_2(\sum_{i_2\in H_2}\xi_{i_2jk}/n_2 - \mu_{2jk})^2\Big],\end{aligned}$$

which implies that for every $j \in T$ and $k \leq s_n$,

$$\begin{aligned}\hat{\omega}_{jk}\omega_{jk}^{-1} - 1 =&(n-2)^{-1}n_1\big[n_1^{-1}\sum_{i\in H_1}\{\omega_{jk}^{-1/2}(\xi_{ijk} - \mu_{1jk})\}^2 - 1\big]\\ &+ (n-2)^{-1}n_2\big[n_2^{-1}\sum_{i'\in H_2}\{\omega_{jk}^{-1/2}(\xi_{i'jk} - \mu_{2jk})\}^2 - 1\big]\\ &- (n-2)^{-1}n_1\big[n_1^{-1}\sum_{i_1\in H_1}\omega_{jk}^{-1/2}(\xi_{i_1jk} - \mu_{1jk})\big]^2\\ &- (n-2)^{-1}n_2\big[n_2^{-1}\sum_{i_2\in H_2}\omega_{jk}^{-1/2}(\xi_{i_2jk} - \mu_{2jk})\big]^2 + 2(n-2)^{-1}.\end{aligned}$$

Together with (66), we obtain

$$\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le 2n^{-1}n_1\Upsilon_1 + 2n^{-1}n_2\Upsilon_2 + 2n^{-1}n_1\Upsilon_3^2 + 2n^{-1}n_2\Upsilon_4^2 + 3n^{-1}, \quad (67)$$

where

$$\Upsilon_1 = \max_{j\in T}\max_{k\le s_n}\big|n_1^{-1}\sum_{i\in H_1}\{\omega_{jk}^{-1/2}(\xi_{ijk}-\mu_{1jk})\}^2 - 1\big|,$$

$$\Upsilon_2 = \max_{j\in T}\max_{k\le s_n}\big|n_2^{-1}\sum_{i'\in H_2}\{\omega_{jk}^{-1/2}(\xi_{i'jk}-\mu_{2jk})\}^2 - 1\big|,$$

$$\Upsilon_3 = \max_{j\in T}\max_{k\le s_n}\big|n_1^{-1}\sum_{i_1\in H_1}\omega_{jk}^{-1/2}(\xi_{i_1jk}-\mu_{1jk})\big|,$$

$$\Upsilon_4 = \max_{j\in T}\max_{k\le s_n}\big|n_2^{-1}\sum_{i_2\in H_2}\omega_{jk}^{-1/2}(\xi_{i_2jk}-\mu_{2jk})\big|.$$

At this point, note that for every $i \in H_1, j \le q_n, k \le s_n$, the sub-exponential norms of the sub-exponential random variables $\{\omega_{jk}^{-1/2}(\xi_{ijk}-\mu_{1jk})\}^2$ satisfy

$$\|\{\omega_{jk}^{-1/2}(\xi_{ijk}-\mu_{1jk})\}^2\|_\psi \le \max\{4\pi, 2e^{2/e}\}. \quad (68)$$

For the term $\Upsilon_1$, conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, one can show that for any $t \ge 0$,

$$\begin{aligned}
&P\big[\Upsilon_1 \ge t\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\le&\sum_{j\in T}\sum_{k\le s_n} P\Big[\big|n_1^{-1}\sum_{i\in H_1}\{\omega_{jk}^{-1/2}(\xi_{ijk}-\mu_{1jk})\}^2 - 1\big| \ge t\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big] \\
\le&2q_n s_n \exp\big[-c_1\min\{t^2, t\}n\big],
\end{aligned} \quad (69)$$

for some universal constant $c_1 > 0$, where the first inequality holds from the union bound inequality, and the second inequality follows from (68) and the Bernstein inequality in Lemma H.2 of Ning and Liu (2017). Similar reasoning gives the result that for any $t \ge 0$,

$$P\big[\Upsilon_2 \ge t\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \le 2q_n s_n \exp\big[-c_2\min\{t^2, t\}n\big], \quad (70)$$

for some universal constant $c_2 > 0$. Regarding the term $\Upsilon_3$, it is clear that for any $t \geq 0$,

$$
\begin{aligned}
&P\big[\Upsilon_3 \geq t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\leq &\sum_{j \in T} \sum_{k \leq s_n} P\Big[\big|n_1^{-1} \sum_{i_1 \in H_1} \omega_{jk}^{-1/2}(\xi_{i_1 jk} - \mu_{1jk})\big| \geq t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big] \\
\leq &2 q_n s_n \exp(-c_3 n t^2), \qquad (71)
\end{aligned}
$$

for some universal constant $c_3 > 0$, where the first inequality is based on the union bound inequality, and the second inequality follows from Hoeffding inequality. Similar argument leads to the result that for any $t \geq 0$,

$$
P\big[\Upsilon_4 \geq t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \leq 2 q_n s_n \exp(-c_4 n t^2), \qquad (72)
$$

for some universal constant $c_4 > 0$. To this end, conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, it can be deduced that for any $t \geq 0$,

$$
\begin{aligned}
&P\big[\|\hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \geq t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\leq &P\big[2n^{-1} n_1 \Upsilon_1 + 2n^{-1} n_2 \Upsilon_2 + 2n^{-1} n_1 \Upsilon_3^2 + 2n^{-1} n_2 \Upsilon_4^2 + 3n^{-1} \geq t \\
&\quad \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\leq &P\big[\Upsilon_1 + \Upsilon_2 + \Upsilon_3^2 + \Upsilon_4^2 + n^{-1} \geq c_5 t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\leq &P\big[\Upsilon_1 \geq 5^{-1} c_5 t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] + P\big[\Upsilon_2 \geq 5^{-1} c_5 t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
&\quad + P\big[\Upsilon_3 \geq 5^{-1/2} c_5^{1/2} t^{1/2} \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
&\quad + P\big[\Upsilon_4 \geq 5^{-1/2} c_5^{1/2} t^{1/2} \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
&\quad + P\big[n^{-1} \geq 5^{-1} c_5 t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\leq &4 q_n s_n \exp\big[-c_6 \min\{t^2, t\} n\big] + 4 q_n s_n \exp(-c_6 n t) + P(n^{-1} \geq 5^{-1} c_5 t) \\
\leq &8 q_n s_n \exp\big[-c_6 \min\{t^2, t\} n\big] + P(n^{-1} \geq 5^{-1} c_5 t),
\end{aligned}
$$

for some carefully chosen universal constants $c_5 > 0$ and $c_6 > 0$, where the first inequality is by (67), the second inequality comes from the definition of $\mathcal{M}_n$ in Lemma 3, the fourth inequality is based on (69), (70), (71) and (72). Accordingly, we set $c_7 = (2c_6^{-1})^{1/2}$ and substitute $t = c_7\{\log(q_n s_n)/n\}^{1/2}$ into the above inequality to obtain

$$P\big[\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le c_7\{\log(q_n s_n)/n\}^{1/2}\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \ge 1 - 8(q_n s_n)^{-1}. \tag{73}$$

It then follows that

$$\begin{aligned}
&P\big[\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le c_7\{\log(q_n s_n)/n\}^{1/2}\big] \\
\ge &\sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le c_7\{\log(q_n s_n)/n\}^{1/2}\big|\{Y_i = y_i\}_{i=1}^n\big] \cdot P\big[\{Y_i = y_i\}_{i=1}^n\big] \\
\ge &\{1 - 8(q_n s_n)^{-1}\} \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[\{Y_i = y_i\}_{i=1}^n\big] = \{1 - 8(q_n s_n)^{-1}\} P(\mathcal{M}_n) \\
\ge &1 - 8\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\},
\end{aligned}$$

where the second inequality is by (73), and the last inequality follows from Lemma 3. Therefore, property 1) holds from the above inequality. Moreover, it can be verified that under the event $\big\{\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le c_7\{\log(q_n s_n)/n\}^{1/2}\big\}$,

$$\|\Lambda_T^{(1)}\hat{\Lambda}_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le 2\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max}.$$

Hence, based on the above two inequalities, we conclude that

$$\begin{aligned}
&P\big[\|\Lambda_T^{(1)}\hat{\Lambda}_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le 2c_7\{\log(q_n s_n)/n\}^{1/2}\big] \\
\ge &1 - 8\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\},
\end{aligned} \tag{74}$$

which completes the proof of property 2). Property 3) can be directly proved by using the fact that $\|\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} - I_{q_n s_n}\|_{\max} \le \|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max}$. Likewise, one can show property 4), which finishes the proof. □

**Lemma 6.** *Assume the following conditions (a)–(b):*

*(a)* $\sup_{j\le p_n}\sum_{k=1}^{\infty}\omega_{jk}<\infty$, $\lambda_{\min}(\Lambda_N^{(1)})\ge c_0 s_n^{-a}$ *for some constants* $c_0>0$ *and* $a>1$.

*(b)* $s_n^{2a}\log\{(p_n-q_n)s_n\}=o(n)$.

*Then there exist universal constants* $c_1>0$ *and* $c_2>0$ *such that:*

*1)* $P\big(\|\hat{\Lambda}_N^{(1)}\Lambda_N^{(1)-1}-I_{(p_n-q_n)s_n}\|_{\max}\le c_1[\log\{(p_n-q_n)s_n\}/n]^{1/2}\big)\ge 1-c_2[\{(p_n-q_n)s_n\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)]$.

*2)* $P\big(\|\Lambda_N^{(1)}\hat{\Lambda}_N^{(1)-1}-I_{(p_n-q_n)s_n}\|_{\max}\le c_1[\log\{(p_n-q_n)s_n\}/n]^{1/2}\big)\ge 1-c_2[\{(p_n-q_n)s_n\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)]$.

*3)* $P\big(\|\hat{\Lambda}_N^{(1)1/2}\Lambda_N^{(1)-1/2}-I_{(p_n-q_n)s_n}\|_{\max}\le c_1[\log\{(p_n-q_n)s_n\}/n]^{1/2}\big)\ge 1-c_2[\{(p_n-q_n)s_n\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)]$.

*4)* $P\big(\|\Lambda_N^{(1)1/2}\hat{\Lambda}_N^{(1)-1/2}-I_{(p_n-q_n)s_n}\|_{\max}\le c_1[\log\{(p_n-q_n)s_n\}/n]^{1/2}\big)\ge 1-c_2[\{(p_n-q_n)s_n\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)]$.

*5)* $P\{det(\hat{\Lambda}_N^{(1)})\neq 0\}\ge 1-c_2[\{(p_n-q_n)s_n\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)]$.

*Note that* $I_{(p_n-q_n)s_n}$ *denotes the* $(p_n-q_n)s_n\times(p_n-q_n)s_n$ *identity matrix.*

*Proof of Lemma 6*: The proof of property 1) is analogous to that of property 1) in Lemma 5. Then, it can be deduced that there exists $c_3>0$ and $c_4>0$ such that with probability at least $1-c_3[\{(p_n-q_n)s_n\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)]$,

$$\lambda_{min}(\hat{\Lambda}_N^{(1)})\ge\lambda_{min}(\Lambda_N^{(1)})-\lambda_{max}(\Lambda_N^{(1)})\|\hat{\Lambda}_N^{(1)}\Lambda_N^{(1)-1}-I_{(p_n-q_n)s_n}\|_{\max}\ge c_4 s_n^{-a},$$

where the last inequality is based on (a), (b) and property 1). As a result, property 5) holds true from the above inequality. Finally, properties 2) to 4) can be derived in a similar fashion as properties 2) to 4) in Lemma 5, which finishes the proof. □

**Lemma 7.** *Assume the following condition (a):*

*(a)* $\log(q_n s_n) = o(n)$.

*Then there exist universal constants* $c_1 > 0$ *and* $c_2 > 0$ *such that:*

$$P\big[\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2 \le c_1\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2}\big]$$

$$\ge 1 - c_2\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}.$$

*Proof of Lemma 7*: First of all, we note that

$$\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2 \le \Omega_1 + \Omega_2 + \Omega_3 + \Omega_4 + \Omega_5, \tag{75}$$

where

$$\begin{aligned}
\Omega_1 =& 2n^{-1} n_1 q_n s_n \max_{j_2 \in T} \max_{k_2 \le s_n} \max_{j_1 \in T} \max_{k_1 \le s_n} \Big| n_1^{-1} \sum_{i \in H_1} \Big[ \big\{\omega_{j_1 k_1}^{-1/2}(\xi_{i j_1 k_1} - \mu_{1 j_1 k_1})\big\} \\
& \cdot \big\{\omega_{j_2 k_2}^{-1/2}(\xi_{i j_2 k_2} - \mu_{1 j_2 k_2})\big\} - \operatorname{corr}(\xi_{j_1 k_1}, \xi_{j_2 k_2}) \Big] \Big|, \\
\Omega_2 =& 2n^{-1} n_2 q_n s_n \max_{j_2 \in T} \max_{k_2 \le s_n} \max_{j_1 \in T} \max_{k_1 \le s_n} \Big| n_2^{-1} \sum_{i \in H_2} \Big[ \big\{\omega_{j_1 k_1}^{-1/2}(\xi_{i j_1 k_1} - \mu_{2 j_1 k_1})\big\} \\
& \cdot \big\{\omega_{j_2 k_2}^{-1/2}(\xi_{i j_2 k_2} - \mu_{2 j_2 k_2})\big\} - \operatorname{corr}(\xi_{j_1 k_1}, \xi_{j_2 k_2}) \Big] \Big|, \\
\Omega_3 =& 2n^{-1} n_1 q_n s_n \max_{j_2 \in T} \max_{k_2 \le s_n} \max_{j_1 \in T} \max_{k_1 \le s_n} \Big| \big\{ n_1^{-1} \sum_{i_1 \in H_1} \omega_{j_1 k_1}^{-1/2}(\xi_{i_1 j_1 k_1} - \mu_{1 j_1 k_1}) \big\} \\
& \cdot \big\{ n_1^{-1} \sum_{i_1 \in H_1} \omega_{j_2 k_2}^{-1/2}(\xi_{i_1 j_2 k_2} - \mu_{1 j_2 k_2}) \big\} \Big|, \\
\Omega_4 =& 2n^{-1} n_2 q_n s_n \max_{j_2 \in T} \max_{k_2 \le s_n} \max_{j_1 \in T} \max_{k_1 \le s_n} \Big| \big\{ n_2^{-1} \sum_{i_2 \in H_2} \omega_{j_1 k_1}^{-1/2}(\xi_{i_2 j_1 k_1} - \mu_{2 j_1 k_1}) \big\} \\
& \cdot \big\{ n_2^{-1} \sum_{i_2 \in H_2} \omega_{j_2 k_2}^{-1/2}(\xi_{i_2 j_2 k_2} - \mu_{2 j_2 k_2}) \big\} \Big|, \\
\Omega_5 =& 4n^{-1} q_n s_n.
\end{aligned}$$

For the term $\Omega_1$, conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, it can be shown that for any $t \geq 0$,

$$
\begin{aligned}
&P\big[\Omega_1 \geq t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\leq &P\Big(\max_{j_2 \in T} \max_{k_2 \leq s_n} \max_{j_1 \in T} \max_{k_1 \leq s_n} \Big| n_1^{-1} \sum_{i \in H_1} \Big[\big\{\omega_{j_1 k_1}^{-1/2}(\xi_{ij_1k_1} - \mu_{1j_1k_1})\big\} \cdot \big\{\omega_{j_2 k_2}^{-1/2}(\xi_{ij_2k_2} - \mu_{1j_2k_2})\big\} \\
&\quad - \mathrm{corr}(\xi_{j_1k_1}, \xi_{j_2k_2})\Big]\Big| \geq (3\pi_1 q_n s_n)^{-1} t \Big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big) \\
\leq &\sum_{j_2 \in T} \sum_{k_2=1}^{s_n} \sum_{j_1 \in T} \sum_{k_1=1}^{s_n} P\Big(\Big| n_1^{-1} \sum_{i \in H_1} \Big[\big\{\omega_{j_1 k_1}^{-1/2}(\xi_{ij_1k_1} - \mu_{1j_1k_1})\big\} \cdot \big\{\omega_{j_2 k_2}^{-1/2}(\xi_{ij_2k_2} - \mu_{1j_2k_2})\big\} \\
&\quad - \mathrm{corr}(\xi_{j_1k_1}, \xi_{j_2k_2})\Big]\Big| \geq (3\pi_1 q_n s_n)^{-1} t \Big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big) \\
\leq &\sum_{j_2 \in T} \sum_{k_2=1}^{s_n} \sum_{j_1 \in T} \sum_{k_1=1}^{s_n} 2 \exp\big[-c_1 n \min\{(q_n s_n)^{-2} t^2, (q_n s_n)^{-1} t\}\big] \\
= &2(q_n s_n)^2 \exp\big[-c_1 n \min\{(q_n s_n)^{-2} t^2, (q_n s_n)^{-1} t\}\big],
\end{aligned}
$$

for some universal constant $c_1 > 0$, where the first inequality is by the definition of $\mathcal{M}_n$ in Lemma 3, the second inequality holds from the union bound inequality, and the last inequality is based on Bernstein inequality and the definition of $\mathcal{M}_n$. To this end, we set $c_2 = (c_1/3)^{-1/2}$ and substitute $t = c_2\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2}$ into the above inequality to obtain

$$P\big[\Omega_1 \geq c_2\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2} \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \leq 2(q_n s_n)^{-1}. \tag{76}$$

Similar reasoning yields that

$$P\big[\Omega_2 \geq c_3\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2} \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \leq 2(q_n s_n)^{-1}, \tag{77}$$

for some universal constant $c_3 > 0$. For the term $\Omega_3$, it is apparent to see that for any

$t \geq 0$,

$$
\begin{aligned}
&P\big[\Omega_3 \geq t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\leq &\sum_{j_2 \in T} \sum_{k_2=1}^{s_n} \sum_{j_1 \in T} \sum_{k_1=1}^{s_n} P\Big(\Big| n_1^{-1} \sum_{i_1 \in H_1} \omega_{j_1 k_1}^{-1/2} (\xi_{i_1 j_1 k_1} - \mu_{1 j_1 k_1})\Big| \geq (3\pi_1 q_n s_n)^{-1/2} t^{1/2} \Big| \\
&\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big) + \sum_{j_2 \in T} \sum_{k_2=1}^{s_n} \sum_{j_1 \in T} \sum_{k_1=1}^{s_n} P\Big(\Big| n_1^{-1} \sum_{i_1 \in H_1} \omega_{j_2 k_2}^{-1/2} (\xi_{i_1 j_2 k_2} - \mu_{1 j_2 k_2})\Big| \\
&\geq (3\pi_1 q_n s_n)^{-1/2} t^{1/2} \Big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big) \\
\leq &4(q_n s_n)^2 \exp(-c_4 n q_n^{-1} s_n^{-1} t),
\end{aligned}
$$

for some universal constant $c_4 > 0$, where the last inequality follows from Hoeffding inequality and the definition of $\mathcal{M}_n$. Therefore, we set $c_5 = 3c_4^{-1}$ and plug $t = c_5 q_n s_n \log(q_n s_n)/n$ into the above inequality to obtain

$$
P\big[\Omega_3 \geq c_5 q_n s_n \log(q_n s_n)/n \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \leq 4(q_n s_n)^{-1}.
$$

Similar reasoning leads to

$$
P\big[\Omega_4 \geq c_6 q_n s_n \log(q_n s_n)/n \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \leq 4(q_n s_n)^{-1},
$$

for some universal constant $c_6 > 0$. Accordingly, we set $c_7 = c_2 + c_3 + c_5 + c_6 + 1$. By combining the above two inequalities with (76), (77), and (75), it can be deduced that

$$
\begin{aligned}
&P\big[\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2 \leq c_7 \{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2} \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \\
\geq &1 - 12(q_n s_n)^{-1}. \qquad (78)
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
&P\big[\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}\|_2 \le c_7\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2}\big] \\
\ge & \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}\|_2 \le \\
& c_7 q_n s_n \{\log(q_n s_n)/n\}^{1/2} \big| \{Y_i = y_i\}_{i=1}^n\big] \cdot P\big[\{Y_i = y_i\}_{i=1}^n\big] \\
\ge & \{1 - 12(q_n s_n)^{-1}\} \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[\{Y_i = y_i\}_{i=1}^n\big] = \{1 - 12(q_n s_n)^{-1}\} P(\mathcal{M}_n) \\
\ge & 1 - 12\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\},
\end{aligned}
$$

where the second inequality is by (78), and the last inequality follows from Lemma 3. This completes the proof. □

**Lemma 8.** *Assume the following conditions (a)–(b):*

*(a)* $q_n^2 s_n^2 \log(q_n s_n) = o(n)$.

*(b)* $c_1 \le \lambda_{\min}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \le \lambda_{\max}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \le c_2$, *for some universal constants* $0 < c_1 < c_2$.

*Then we have the following properties:*

*1) There exist universal constants* $c_3 > 0$ *and* $c_4 > 0$ *such that*

$$P(\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2 \le c_3) \ge 1 - c_4\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}.$$

*2) There exist universal constants* $c_5 > 0$ *and* $c_6 > 0$ *such that*

$$P(\|\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}\|_2 \le c_5) \ge 1 - c_6\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}.$$

*Proof of Lemma 8*: First of all, we note that

$$\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2 \le \|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}\|_2 + c_2,$$

where $c_2$ is defined in condition (b). Together with condition (a) and Lemma 7, it can be concluded that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that with probability at least $1 - c_3\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$,

$$\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2 \le c_4\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2} + c_2 \le 2c_2,$$

which completes the proof of property 1). To show the second property, we first notice that

$$\|\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}\|_2 = \lambda_{\min}^{-1}(\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}). \tag{79}$$

Moreover, it is apparent to deduce that

$$\lambda_{\min}(\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \ge c_1 - \|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2,$$

where $c_1$ is defined in condition (b). Together with condition (a) and Lemma 7, we conclude that there exist universal constants $c_5 > 0$ and $c_6 > 0$ such that with probability at least $1 - c_5\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$,

$$\lambda_{\min}(\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \ge c_1 - c_6\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2} \ge c_1/2.$$

Together with (79), the proof is finished. □

**Lemma 9.** *Assume the following conditions (a)–(b):*

*(a)* $q_n^2 s_n^2 \log(q_n s_n) = o(n)$.

*(b)* $c_1 \le \lambda_{\min}(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \le \lambda_{\max}(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \le c_2$, *for some universal constants* $0 < c_1 < c_2$.

*Then there exist universal constants* $c_3 > 0$ *and* $c_4 > 0$ *such that:*

*1)* $$P\Big(\big|\{sgn(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} sgn(\beta_T^{(1)})\}/\{sgn(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2} sgn(\beta_T^{(1)})\} - 1\big| \le c_3\big[\{\log(q_n s_n)/n\}^{1/2} + \{\log\log(n)/n\}^{1/2}\big]\Big)$$
$$\ge 1 - c_4\big[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big].$$

*2)* $P\Big(\big|\{sgn(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\}/\{sgn(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}sgn(\beta_T^{(1)})\}$
$-1\big| \le c_3\big[\{\log(q_ns_n)/n\}^{1/2} + \{\log\log(n)/n\}^{1/2}\big]\Big)$
$\ge 1 - c_4\big[(q_ns_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big].$

*Proof of Lemma 9*: First of all, we note that

$$\text{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}) = \text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}) + \Omega_1 + 2\Omega_2, \tag{80}$$

where

$$\Omega_1 = \text{sgn}(\beta_T^{(1)})'(\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} - I_{q_ns_n})(\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2})(\Lambda_T^{(1)-1/2}\hat{\Lambda}_T^{(1)1/2} - I_{q_ns_n})\text{sgn}(\beta_T^{(1)}),$$
$$\Omega_2 = \text{sgn}(\beta_T^{(1)})'(\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2})(\Lambda_T^{(1)-1/2}\hat{\Lambda}_T^{(1)1/2} - I_{q_ns_n})\text{sgn}(\beta_T^{(1)}).$$

For the term $\Omega_1$, it can be deduced that

$$\Omega_1 \le q_ns_n\|\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\|_2 \cdot \|\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} - I_{q_ns_n}\|_{\max}^2.$$

Together with condition (a), condition (b), Lemma 8, and Lemma 5, it can be concluded that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that

$$P\{\Omega_1 \le c_3q_ns_n\log(q_ns_n)/n\} \ge 1 - c_4\{(q_ns_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}. \tag{81}$$

For the term $\Omega_2$, one has

$$\begin{aligned}|\Omega_2| \le& \|(\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\text{sgn}(\beta_T^{(1)})\|_1 \cdot \|(\Lambda_T^{(1)-1/2}\hat{\Lambda}_T^{(1)1/2} - I_{q_ns_n})\text{sgn}(\beta_T^{(1)})\|_\infty \\ \le& q_ns_n\|\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\|_2 \cdot \|\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} - I_{q_ns_n}\|_{\max}.\end{aligned}$$

Together with condition (a), condition (b), Lemma 8, and Lemma 5, it can be deduced that there exist universal constants $c_5 > 0$ and $c_6 > 0$ such that

$$P[|\Omega_2| \le c_5\{q_n^2s_n^2\log(q_ns_n)/n\}^{1/2}] \ge 1 - c_6\{(q_ns_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}.$$

Together with (80) and (81), it can be concluded that there exist universal constants $c_7 > 0$ and $c_8 > 0$ such that with probability at least $1 - c_7\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}$,

$$
\begin{aligned}
&|\text{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}) - \text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})| \\
\le & c_8\{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2}.
\end{aligned}
$$

Moreover, we note that

$$
\begin{aligned}
&\big|\{\text{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\}/\{\text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\} - 1\big| \\
\le & \{\text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\}^{-1} \\
&\cdot \big|\text{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}) - \text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\big| + \\
&\big|\{\text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\}/\{\text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\} - 1\big| \\
\le & c_2 (q_n s_n)^{-1} \big|\text{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}) - \text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\big| \\
&+ \big|\{\text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\}/\{\text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\} - 1\big|,
\end{aligned}
$$

where the last inequality is based on condition (b). Therefore, by combining Lemma 22 with the above two inequalities, we conclude that there exist universal constants $c_9 > 0$ and $c_{10} > 0$ such that with probability at least $1 - c_9\big[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big]$,

$$
\begin{aligned}
&\big|\{\text{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\}/\{\text{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\} - 1\big| \\
\le & c_{10}\big[\{\log(q_n s_n)/n\}^{1/2} + q_n s_n/n + \{\log\log(n)/n\}^{1/2}\big] \\
\le & 2c_{10}\big[\{\log(q_n s_n)/n\}^{1/2} + \{\log\log(n)/n\}^{1/2}\big],
\end{aligned}
$$

which completes the proof of property 1). To show the second property, we notice the fact that

$$
\begin{aligned}
&\big|\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}/\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\} - 1\big| \\
=&\big|\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}/\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\} - 1\big| \\
&\cdot\big|\{\mathrm{sgn}(\beta_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}/\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\big|^{-1}.
\end{aligned}
$$

Together with property 1), property 2) follows directly, which finishes the proof. □

**Lemma 10.** *Assume the following conditions (a)–(b):*

*(a)* $q_n s_n = o(n)$.

*(b)* $c_1 \le \lambda_{\min}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \le \lambda_{\max}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \le c_2$*, for some universal constants* $0 < c_1 < c_2$.

*Then there exist universal constants* $c_3 > 0$ *and* $c_4 > 0$ *such that with probability at least* $1 - c_3[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$*, we have:*

$$
\begin{aligned}
&|\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}sgn(\beta_T^{(1)}) - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})| \\
\le& c_4[q_n s_n/n + \{\log(q_n s_n)/n\}^{1/2} + \{\log\log(n)/n\}^{1/2}] \cdot |\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})| \\
&+ c_4(q_n s_n)^{1/2}\{\log\log(n)/n\}^{1/2}[1 + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log\log(n)/n\}^{1/2}]^{1/2} \\
&+ c_4(q_n s_n)^{1/2}\{\log(q_n s_n)/n\}^{1/2}\{q_n s_n\log(q_n s_n\log n)/n\}^{1/2} \\
&\cdot[1 + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log\log(n)/n\}^{1/2}]^{1/2}.
\end{aligned}
$$

*Proof of Lemma 10*: First of all, we note that

$$
|\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| \le \Omega_1 + \Omega_2, \tag{82}
$$

where

$$
\begin{aligned}
\Omega_1 &= |\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})|, \\
\Omega_2 &= |\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) - \hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})|.
\end{aligned}
$$

For the term $\Omega_1$, Lemma 21 together with condition (b) imply that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that with probability at least $1 - c_3[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\Omega_1 \leq c_4\{q_n s_n \log\log(n)/n\}^{1/2}. \tag{83}$$

For the term $\Omega_2$, it is clear that

$$\Omega_2 \leq \Pi_1 + \Pi_2, \tag{84}$$

where

$$\Pi_1 = |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} (\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} - I_{q_n s_n}) \mathrm{sgn}(\beta_T^{(1)})|,$$
$$\Pi_2 = |\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)}) - \hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})|.$$

For the term $\Pi_1$, it is not difficult to verify that

$$\begin{aligned}\Pi_1 \leq & \|\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} - I_{q_n s_n}\|_{\max} \cdot \{|\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| \\ & + \|\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2}\|_1\}.\end{aligned}$$

To bound the term $\|\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2}\|_1$, based on Lemma 23, and conditions (a) and (b), it can be deduced that there exist universal constants $c_5 > 0$ and $c_6 > 0$ such that with probability at least $1 - c_5[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\begin{aligned}& \|\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2}\|_1 \\ \leq & c_6[q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}] \cdot |\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| + c_6 q_n s_n \\ & \cdot \{\log(q_n s_n \log n)/n\}^{1/2} [1 + \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} + \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2}.\end{aligned}$$

To this end, by combining the above two inequalities with Lemma 5, it can be concluded that there exist universal constants $c_7 > 0$ and $c_8 > 0$ such that with probability at least

$1 - c_7[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\begin{aligned}\Pi_1 \leq & c_8\{\log(q_n s_n)/n\}^{1/2} \cdot |\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| + \\ & c_8\{q_n s_n \log(q_n s_n)/n\}^{1/2}\{q_n s_n \log(q_n s_n \log n)/n\}^{1/2} \\ & \cdot [1 + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2}. \end{aligned} \tag{85}$$

To bound the term $\Pi_2$, we note that

$$\Pi_2 \leq c_1^{-1} q_n s_n (1 + \Upsilon_1) \cdot |\Upsilon_2| + \{|\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| + \Upsilon_3\} \cdot \Upsilon_1, \tag{86}$$

where $c_1$ is defined in condition (b) and

$$\begin{aligned}\Upsilon_1 =& |\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1} - 1|, \\ \Upsilon_2 =& \{\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1} \\ & - \{\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1}, \\ \Upsilon_3 =& |\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})|. \end{aligned}$$

For the term $\Upsilon_1$, Lemma 22 entails that there exist universal constants $c_9 > 0$ and $c_{10} > 0$ such that with probability at least $1 - c_9[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\Upsilon_1 \leq c_{10}[q_n s_n/n + \{\log\log(n)/n\}^{1/2}]. \tag{87}$$

For the term $\Upsilon_2$, by using similar arguments as in the proof of Lemma 23, it can be deduced that there exist universal constants $c_{11} > 0$ and $c_{12} > 0$ such that conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}$, and for any $t \geq 0$,

$$\begin{aligned} & P\big[|\Upsilon_2| \geq t \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}\big] \\ \leq & c_{11} \exp\big[-c_{12} n\{\hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}^{-1}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}t^2\big]. \end{aligned}$$

By plugging $t = c_{13}\{\hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2}\{\log\log(n)/n\}^{1/2}$ with $c_{13} = c_{12}^{-1/2}$ into the above inequality, it yields that

$$
\begin{aligned}
&P\big[|\Upsilon_2| \le c_{13}\{\hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\qquad \cdot \{\log\log(n)/n\}^{1/2}\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}\big] \\
&\ge 1 - c_{11}\{\log(n)\}^{-1}. \qquad (88)
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&P\big[|\Upsilon_2| \le c_{13}\{\hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\quad \cdot \{\log\log(n)/n\}^{1/2}\big] \\
&\ge \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[|\Upsilon_2| \le c_{13}\{\hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\quad \cdot \{\log\log(n)/n\}^{1/2}\big|\{Y_i = y_i\}_{i=1}^n\big] \cdot P\big[\{Y_i = y_i\}_{i=1}^n\big] \\
&= \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} \Big\{\int_{\hat{\nu}_T} P\big[|\Upsilon_2| \le c_{13}\{\hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\quad \cdot \{\log\log(n)/n\}^{1/2}\big|\{Y_i = y_i\}_{i=1}^n \cap \{\hat{\nu}_T\}\big] \cdot f(\hat{\nu}_T|\{Y_i = y_i\}_{i=1}^n)d\hat{\nu}_T\Big\} \cdot P\big[\{Y_i = y_i\}_{i=1}^n\big] \\
&\ge [1 - c_{11}\{\log(n)\}^{-1}] \cdot \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[\{Y_i = y_i\}_{i=1}^n\big] = [1 - c_{11}\{\log(n)\}^{-1}] \cdot P(\mathcal{M}_n) \\
&\ge 1 - c_{14}\big[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big],
\end{aligned}
$$

for some universal constant $c_{14} > 0$, where $f(\hat{\nu}_T|\{Y_i = y_i\}_{i=1}^n)$ denotes the conditional density function, and the second inequality is by (88). Together with Lemma 19 yields the result that there exist universal constants $c_{15} > 0$ and $c_{16} > 0$ such that with probability at least $1 - c_{15}[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$
\begin{aligned}
|\Upsilon_2| \le & c_{16}[q_n s_n/n + \log\log(n)/n + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log\log(n)/n\}^{1/2}]^{1/2} \\
&\cdot (q_n s_n)^{-1/2}\{\log\log(n)/n\}^{1/2}. \qquad (89)
\end{aligned}
$$

For the term $\Upsilon_3$, Lemma 21 leads to the result that there exist universal constants $c_{17} > 0$ and $c_{18} > 0$ such that with probability at least $1 - c_{17}[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\Upsilon_3 \leq c_{18}\{q_n s_n \log\log(n)/n\}^{1/2}.$$

Together with (87), (89) and (86), it can be observed that there exist universal constants $c_{19} > 0$ and $c_{20} > 0$ such that with probability at least $1 - c_{19}[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\begin{aligned}
\Pi_2 \leq\ & c_{20}[q_n s_n/n + \log\log(n)/n + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2} \\
& \cdot (q_n s_n)^{1/2}\{\log\log(n)/n\}^{1/2} + c_{20}[q_n s_n/n + \{\log\log(n)/n\}^{1/2}] \cdot |\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})| \\
& + c_{20}\{q_n s_n \log\log(n)/n\}^{1/2} \cdot [q_n s_n/n + \{\log\log(n)/n\}^{1/2}].
\end{aligned}$$

Together with (84) and (85), there exist universal constants $c_{21} > 0$ and $c_{22} > 0$ such that with probability at least $1 - c_{21}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\begin{aligned}
\Omega_2 \leq\ & c_{22}(q_n s_n)^{1/2}\{\log\log(n)/n\}^{1/2} \\
& \cdot [q_n s_n/n + \log\log(n)/n + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2} \\
& + c_{22}[q_n s_n/n + \{\log(q_n s_n)/n\}^{1/2} + \{\log\log(n)/n\}^{1/2}] \cdot |\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})| \\
& + c_{22}\{q_n s_n \log\log(n)/n\}^{1/2} \cdot [q_n s_n/n + \{\log\log(n)/n\}^{1/2}] \\
& + c_{22}\{q_n s_n \log(q_n s_n)/n\}^{1/2}\{q_n s_n \log(q_n s_n \log n)/n\}^{1/2} \\
& \cdot [1 + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2}.
\end{aligned}$$

Together with (82) and (83), it can be concluded that there exist universal constants $c_{23} > 0$ and $c_{24} > 0$ such that with probability at least $1 - c_{23}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} +$

$\exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$|\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)}) - \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})|$$

$$\leq c_{24}[q_n s_n/n + \{\log(q_n s_n)/n\}^{1/2} + \{\log\log(n)/n\}^{1/2}] \cdot |\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})|$$

$$+ c_{24}(q_n s_n)^{1/2} \{\log\log(n)/n\}^{1/2} [1 + \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} + \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2}$$

$$+ c_{24}(q_n s_n)^{1/2} \{\log(q_n s_n)/n\}^{1/2} \{q_n s_n \log(q_n s_n \log n)/n\}^{1/2}$$

$$\cdot [1 + \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} + \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2},$$

which completes the proof. □

**Lemma 11.** *Assume the following conditions (a)–(d):*

*(a)* $\max\{q_n^2 s_n^2 \log(q_n s_n), q_n s_n \log(p_n - q_n)\} = o(n)$.

*(b)* $c_1 \leq \lambda_{\min}(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq \lambda_{\max}(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq c_2$, *for some universal constants* $0 < c_1 < c_2$.

*(c)* $K_1 \log\{(p_n - q_n) s_n \log n\}/(n\lambda_n^2) \leq \sum_{j\in T} \sum_{k=1}^{s_n} \omega_{jk} \beta_{jk}^2 \leq K_2 \log\{(p_n - q_n) s_n \log n\}/(n\lambda_n^2) \to \infty$, *for some sufficiently large universal constants* $K_2 > K_1 > 0$.

*(d)* $\min_{j\in T} \min_{k\leq s_n} \omega_{jk}^{1/2} |\beta_{jk}| > K_3[\log\{(p_n - q_n) s_n \log n\}/(n\lambda_n^2)]^{1/2} \{\log(q_n s_n \log n)/n\}^{1/2} + K_3[\log\{(p_n - q_n) s_n \log n\}/(n\lambda_n)] \cdot \|\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} sgn(\beta_T^{(1)})\|_\infty + K_3[\log\{(p_n - q_n) s_n \log n\}/(n\lambda_n)] \cdot [\{q_n s_n \log(q_n s_n)/n\}^{1/2} + \{q_n s_n \log\log(n)/n\}^{1/2}]$, *for some sufficiently large universal constant* $K_3 > 0$.

*Then there exists a universal constant $c_3 > 0$ such that:*

$$
\begin{aligned}
&P\{sgn(\tilde{v}_T) = sgn(\beta_T^{(1)}) = sgn(\hat{\beta}_T^{(1)})\} \\
\geq &1 - c_3[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)],
\end{aligned}
$$

*where $\hat{\beta}_T^{(1)} = S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}$, and also recall that $\tilde{v}_T$ is defined in Lemma 2.*

*Proof of Lemma 11*: First of all, we denote the two index sets $\mathcal{S}_1$ and $\mathcal{S}_2$ as

$$
\mathcal{S}_1 = \{k : e_k'\Sigma_{TT}^{(1)-1}\nu_T^{(1)} > 0\}, \qquad \mathcal{S}_2 = \{k : e_k'\Sigma_{TT}^{(1)-1}\nu_T^{(1)} < 0\}.
$$

By definition, we have $\mathcal{S}_1 \cup \mathcal{S}_2 = \{1, \ldots, q_n s_n\}$. Moreover, by using Lemma 23 and conditions (a)–(c), it can be shown that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that

$$
\begin{aligned}
&P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} \geq e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \\
&\quad - c_3[q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}] \cdot e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \\
&\quad - c_3\{\log(q_n s_n \log n)/n\}^{1/2} \cdot \{\nu_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{1/2}\Big\}\Big] \\
\geq &1 - c_4[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]. \qquad (90)
\end{aligned}
$$

For the term $\nu_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}$, conditions (b) and (c) entail that

$$
\nu_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \sim \log\{(p_n - q_n)s_n \log n\}/(n\lambda_n^2) \to \infty, \quad \text{as} \quad n \to \infty.
$$

Together with (90), there exist positive universal constants $c_5$, $c_6$ and $c_7$ such that

$$
\begin{aligned}
&P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} \geq c_5 e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \\
&\quad - c_6[\log\{(p_n - q_n)s_n \log n\}/(n\lambda_n^2)]^{1/2}\{\log(q_n s_n \log n)/n\}^{1/2}\Big\}\Big] \\
\geq &1 - c_7[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]. \qquad (91)
\end{aligned}
$$

By choosing $K_3 > c_6/c_5$ in condition (d), (91) together with condition (d) further implies that

$$P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} > 0\Big\}\Big] \geq 1 - c_7[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].$$

Likewise, it can be deduced that there exists a universal constant $c_8 > 0$ such that

$$P\Big[\bigcap_{k\in\mathcal{S}_2}\Big\{e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} < 0\Big\}\Big] \geq 1 - c_8[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].$$

Putting the above two inequalities together implies that there exists a universal constant $c_9 > 0$ such that

$$P\{\mathrm{sgn}(\beta_T^{(1)}) = \mathrm{sgn}(\hat{\beta}_T^{(1)})\} \geq 1 - c_9[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]. \quad (92)$$

Moreover, it can be recalled from Lemma 2 that the quantity $\tilde{v}_T$ can be formulated as

$$\tilde{v}_T = \hat{\vartheta}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \lambda_n S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}),$$

where

$$\begin{aligned}\hat{\vartheta} =& \{n_1n_2n^{-1}(n-2)^{-1}\}\{1 + \lambda_n\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\} \\ &\cdot \big[1 + \{n_1n_2n^{-1}(n-2)^{-1}\}\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\big]^{-1}.\end{aligned}$$

To this end, by combining conditions (a)–(c) with Lemma 10, it can be deduced that there exists a universal constant $c_{10} > 0$ such that with probability at least $1 - c_{10}[(q_ns_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\lambda_n\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) = \lambda_n\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\{1+o(1)\} + o(1).$$

Similarly, by combining conditions (a)–(c) with Lemma 4, it can be deduced that there exists a universal constant $c_{11} > 0$ such that with probability at least $1 - c_{11}[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\hat{\nu}_T^{(1)'}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} = \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\{1+o(1)\}.$$

According to the above three inequalities and Lemma 3, it can be concluded that there exist universal constants $c_{12} > 0$ and $c_{13} > 0$ such that with probability at least $1 - c_{12}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\hat{\vartheta} \geq c_{13}\pi_1\pi_2\{1 + \lambda_n \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\}\{1 + \pi_1\pi_2 \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1}.$$

For the term $\lambda_n \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})$, one has

$$\begin{aligned}
&\lambda_n \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}) \leq \lambda_n \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2} \times \\
&\{\text{sgn}(\beta_T^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)})\}^{1/2} \lesssim \lambda_n \{q_n s_n \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2} \\
&\lesssim [q_n s_n \log\{(p_n - q_n) s_n \log n\}/n]^{1/2} \lesssim o(1),
\end{aligned} \tag{93}$$

where the second and the third inequalities are based on (b) and (c), and the last inequality follows from (a). Piecing the above two inequalities together yields that there exist universal constants $c_{14} > 0$ and $c_{15} > 0$ such that with probability at least $1 - c_{14}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\hat{\vartheta} \geq c_{15}\{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1}.$$

Together with (91) and (92), it can be deduced that there exist universal constants $c_{16}, c_{17}, c_{18} > 0$ such that

$$\begin{aligned}
&P\Big[\bigcap_{k \in \mathcal{S}_1} \Big\{\hat{\vartheta} e_k' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} \geq c_{17}\{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{-1} (e_k' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \\
&\quad - c_{18}[\log\{(p_n - q_n) s_n \log n\}/(n\lambda_n^2)]^{1/2} \{\log(q_n s_n \log n)/n\}^{1/2})\Big\}\Big] \\
&\geq 1 - c_{16}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].
\end{aligned}$$

In addition, utilizing Lemma 24 and conditions (a)–(c), it can also be justified that there exist universal constants $c_{19} > 0$ and $c_{20} > 0$ such that

$$
\begin{aligned}
P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{&\lambda_n|e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| \le \lambda_n|e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| \\
&+ c_{19}\lambda_n[q_ns_n/n + \{\log(q_ns_n\log n)/n\}^{1/2}]\cdot|e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| \\
&+ c_{19}\lambda_n[\{q_ns_n\log(q_ns_n)/n\}^{1/2} + \{q_ns_n\log\log(n)/n\}^{1/2}]\Big\}\Big] \\
\ge& 1 - c_{20}[(q_ns_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].
\end{aligned}
$$

Based on the above two inequalities, it is seen that there exist positive universal constants $c_{21}$, $c_{22}$ and $c_{23}$ that

$$
\begin{aligned}
P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{&e_k'\Lambda_T^{(1)1/2}\tilde{v}_T \ge c_{21}\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\}^{-1} \\
&\big(e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} - c_{22}[\log\{(p_n-q_n)s_n\log n\}/(n\lambda_n^2)]^{1/2}\{\log(q_ns_n\log n)/n\}^{1/2} \\
&- c_{22}[\log\{(p_n-q_n)s_n\log n\}/(n\lambda_n)]\cdot|e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| \\
&- c_{22}[\log\{(p_n-q_n)s_n\log n\}/(n\lambda_n)]\cdot[\{q_ns_n\log(q_ns_n)/n\}^{1/2} + \{q_ns_n\log\log(n)/n\}^{1/2}]\big)\Big\}\Big] \\
\ge& 1 - c_{23}[(q_ns_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].
\end{aligned}
$$

By choosing $K_3 > c_{22}$ in condition (d), it follows from condition (d) and the above inequality that

$$
P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{e_k'\Lambda_T^{(1)1/2}\tilde{v}_T > 0\Big\}\Big] \ge 1 - c_{23}[(q_ns_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].
$$

Similar reasoning leads to the result that there exists a universal constants $c_{24} > 0$ such that

$$
P\Big[\bigcap_{k\in\mathcal{S}_2}\Big\{e_k'\Lambda_T^{(1)1/2}\tilde{v}_T < 0\Big\}\Big] \ge 1 - c_{24}[(q_ns_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].
$$

Based on (92) and the above two inequalities, there exists a universal constant $c_{25} > 0$ such that

$$P\{\text{sgn}(\tilde{v}_T) = \text{sgn}(\beta_T^{(1)}) = \text{sgn}(\hat{\beta}_T^{(1)})\}$$
$$\geq 1 - c_{25}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)],$$

which concludes the proof. □

**Lemma 12.** *Let $a_n$ and $b_n$ be any two sequences of constants such that $a_n \to \infty$ and $b_n \to 0$. Also let $X_n$ and $U_n$ be any two sequences of random variables such that $X_n = o_p(1)$ and $U_n = o_p(1)$. Assume that we have the following conditions (a)–(b):*

*(a)* $a_n X_n = o_p(1)$.

*(b)* $a_n^{1/2}(U_n - b_n) = o_p(1)$.

*Then we have the following property:*

$$\Phi(-a_n^{1/2}(1 + X_n) + U_n)/\Phi(-a_n^{1/2} + b_n) \xrightarrow{p} 1.$$

*Proof of Lemma 12*: The proof is analogous to that of Lemma 1 in Shao et al. (2011). □

**Lemma 13.** *Consider a pair $A$, $B$ of $p \times p$ matrices, assume the following condition (a):*

*(a)* $\lambda_{\min}(A - B) \geq 0$.

*Then we have the following property:*

$$\lambda_{\min}(A) \geq \lambda_{\min}(B), \qquad \lambda_{\max}(A) \geq \lambda_{\max}(B).$$

*Proof of Lemma 13*: First of all, we have

$$\lambda_{\min}(A) \geq \lambda_{\min}(A - B) + \lambda_{\min}(B) \geq \lambda_{\min}(B),$$

where the last inequality is by condition (a). Similarly, we also have

$$\lambda_{\max}(A) \geq \lambda_{\min}(A - B) + \lambda_{\max}(B) \geq \lambda_{\max}(B),$$

where the last inequality is by condition (a) as well, which completes the proof. $\square$

**Lemma 14.** *For any $p \times p$ square matrix $A$, partitioned as*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

*where $A_{11}$ is a $k \times k$ matrix for some positive integer $k < p$, assume we have the following condition (a):*

*(a)* $c_1 \leq \lambda_{\min}(A) \leq \lambda_{\max}(A) \leq c_2$, *for some universal constants* $0 < c_1 < c_2$.

*Then we have the following properties:*

*1)* $c_1 \leq \lambda_{\min}(A_{11} - A_{12}A_{22}^{-1}A_{21}) \leq \lambda_{\max}(A_{11} - A_{12}A_{22}^{-1}A_{21}) \leq c_2,$

$c_1 \leq \lambda_{\min}(A_{22} - A_{21}A_{11}^{-1}A_{12}) \leq \lambda_{\max}(A_{22} - A_{21}A_{11}^{-1}A_{12}) \leq c_2.$

*2)* $\lambda_{\max}(A_{12}A_{22}^{-1}A_{21}) \leq \lambda_{\max}(A_{11}) \leq c_2,$

$\lambda_{\max}(A_{21}A_{11}^{-1}A_{12}) \leq \lambda_{\max}(A_{22}) \leq c_2,$

$\lambda_{\min}(A_{12}A_{22}^{-1}A_{21}) \leq \lambda_{\min}(A_{11}),$

$\lambda_{\min}(A_{21}A_{11}^{-1}A_{12}) \leq \lambda_{\min}(A_{22}).$

*Proof of Lemma 14*: Based on condition (a), we have

$$c_2^{-1} \leq \lambda_{\min}(A^{-1}) \leq \lambda_{\max}(A^{-1}) \leq c_1^{-1},$$

where $A^{-1}$ can be expressed as

$$A^{-1} = \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ -A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}.$$

Hence, we have

$$c_2^{-1} \le \lambda_{\min}((A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}) \le \lambda_{\max}((A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}) \le c_1^{-1},$$
$$c_2^{-1} \le \lambda_{\min}((A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}) \le \lambda_{\max}((A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}) \le c_1^{-1},$$

which implies that

$$c_1 \le \lambda_{\min}(A_{11} - A_{12}A_{22}^{-1}A_{21}) \le \lambda_{\max}(A_{11} - A_{12}A_{22}^{-1}A_{21}) \le c_2,$$
$$c_1 \le \lambda_{\min}(A_{22} - A_{21}A_{11}^{-1}A_{12}) \le \lambda_{\max}(A_{22} - A_{21}A_{11}^{-1}A_{12}) \le c_2,$$

finishing the proof of property 1). Finally, by combining property 1) with Lemma 13, the assertion in property 2) follows immediately, which completes the proof. □

**Lemma 15.** *Let $\{X_1, \ldots, X_{n+m}\}$ be a sample of random vectors in $\mathbb{R}^p$. Denote*

$$S_1 = \sum_{i=1}^{n}(X_i - \bar{X}_1)(X_i - \bar{X}_1)'/(n-1), \quad \bar{X}_1 = \sum_{i=1}^{n} X_i/n,$$
$$S_2 = \sum_{i=n+1}^{n+m}(X_i - \bar{X}_2)(X_i - \bar{X}_2)'/(m-1), \quad \bar{X}_2 = \sum_{i=n+1}^{n+m} X_i/m,$$
$$S = \sum_{i=1}^{n+m}(X_i - \bar{X})(X_i - \bar{X})'/(n+m-2), \quad \bar{X} = \sum_{i=1}^{n+m} X_i/(n+m),$$
$$S_{pool} = \{(n-1)S_1 + (m-1)S_2\}/(n+m-2).$$

*Then we have the following property:*

$$S = S_{pool} + nm(n+m)^{-1}(n+m-2)^{-1}(\bar{X}_1 - \bar{X}_2)(\bar{X}_1 - \bar{X}_2)'.$$

*Proof of Lemma 15*: The term $S$ can be decomposed as $S = I_1 + I_2$ with

$$I_1 = \sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})'/(n+m-2),$$
$$I_2 = \sum_{i=n+1}^{n+m}(X_i - \bar{X})(X_i - \bar{X})'/(n+m-2).$$

For the term $I_1$, one has

$$I_1 = (n-1)(n+m-2)^{-1}S_1 + nm^2(n+m)^{-2}(n+m-2)^{-1}(\bar{X}_1 - \bar{X}_2)(\bar{X}_1 - \bar{X}_2)'.$$

By symmetry, we also have

$$I_2 = (m-1)(n+m-2)^{-1}S_2 + mn^2(n+m)^{-2}(n+m-2)^{-1}(\bar{X}_1 - \bar{X}_2)(\bar{X}_1 - \bar{X}_2)'.$$

Based on the above results, we conclude that $S = S_{pool} + nm(n+m)^{-1}(n+m-2)^{-1}(\bar{X}_1 - \bar{X}_2)(\bar{X}_1 - \bar{X}_2)'$, which finishes the proof. □

**Lemma 16.** *Recall that $T = \{1, \ldots, q_n\}$. Assume the matrix $\Sigma_{TT}^{(1)}$ is invertible and consider the following optimization problem:*

$$\min_{w_T \in \mathbb{R}^{q_n s_n}} \Big[\frac{1}{2} w_T' \big(\Sigma_{TT}^{(1)} + \pi_1\pi_2 \nu_T^{(1)}\nu_T^{(1)'}\big) w_T - \pi_1\pi_2 w_T' \nu_T^{(1)} + \lambda_n (\Lambda_T^{(1)1/2} w_T)' sgn(\beta_T^{(1)})\Big],$$

*where $w_T = (w_1', \ldots, w_{q_n}')'$ with sub-vectors $w_j = (w_{j1}, \ldots, w_{js_n})' \in \mathbb{R}^{s_n}$. Let $\tilde{w}_T$ be the solution of the optimization problem where $\tilde{w}_T = (\tilde{w}_1', \ldots, \tilde{w}_{q_n}')'$ with sub-vectors $\tilde{w}_j = (\tilde{w}_{j1}, \ldots, \tilde{w}_{js_n})' \in \mathbb{R}^{s_n}$, then we have:*

$$\tilde{w}_T = \pi_1\pi_2(1 + \lambda_n \|\Lambda_T^{(1)1/2}\beta_T^{(1)}\|_1)(1 + \pi_1\pi_2 \beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)})^{-1}\beta_T^{(1)} - \lambda_n \Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2} sgn(\beta_T^{(1)}).$$

*Proof of Lemma 16*: First of all, based on first order condition, one has

$$\big(\Sigma_{TT}^{(1)} + \pi_1\pi_2\nu_T^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T = \pi_1\pi_2\nu_T^{(1)} - \lambda_n \Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}). \quad (94)$$

Moreover, according to Sherman-Morrison-Woodbury formula, we have

$$\begin{aligned}\big(\Sigma_{TT}^{(1)} + \pi_1\pi_2\nu_T^{(1)}\nu_T^{(1)'}\big)^{-1} =& \Sigma_{TT}^{(1)-1} - \Sigma_{TT}^{(1)-1}\nu_T^{(1)}(\pi_1^{-1}\pi_2^{-1} + \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{-1}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1} \\ =& \Sigma_{TT}^{(1)-1} - \pi_1\pi_2(1 + \pi_1\pi_2\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{-1}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}.\end{aligned}$$

Finally, by combining the above two equations, we have

$$
\begin{aligned}
\tilde{w}_T =& \big(\Sigma_{TT}^{(1)} + \pi_1\pi_2\nu_T^{(1)}\nu_T^{(1)'}\big)^{-1}\{\pi_1\pi_2\nu_T^{(1)} - \lambda_n\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\\
=& \{\Sigma_{TT}^{(1)-1} - \pi_1\pi_2(1+\pi_1\pi_2\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{-1}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\}\\
&\cdot\{\pi_1\pi_2\nu_T^{(1)} - \lambda_n\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\\
=& \{\Sigma_{TT}^{(1)-1} - \pi_1\pi_2(1+\pi_1\pi_2\beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)})^{-1}\beta_T^{(1)}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\}\\
&\cdot\{\pi_1\pi_2\nu_T^{(1)} - \lambda_n\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\\
=& \{\pi_1\pi_2\Sigma_{TT}^{(1)-1}\nu_T^{(1)} - \pi_1^2\pi_2^2(1+\pi_1\pi_2\beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)})^{-1}\beta_T^{(1)}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\} -\\
&\{\lambda_n\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) - \pi_1\pi_2(1+\pi_1\pi_2\beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)})^{-1}\beta_T^{(1)}\lambda_n\|\Lambda_T^{(1)1/2}\beta_T^{(1)}\|_1\}\\
=& \pi_1\pi_2(1+\lambda_n\|\Lambda_T^{(1)1/2}\beta_T^{(1)}\|_1)(1+\pi_1\pi_2\beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)})^{-1}\beta_T^{(1)} - \lambda_n\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}),
\end{aligned}
$$

which finishes the proof. □

**Lemma 17.** *Consider the following optimization problem:*

$$
\min_{w\in\mathbb{R}^{p_n s_n}}\Big[\frac{1}{2}w'\big(\Sigma^{(1)} + \pi_1\pi_2\nu^{(1)}\nu^{(1)'}\big)w - \pi_1\pi_2 w'\nu^{(1)} + \lambda_n\sum_{j=1}^{p_n}\|\Lambda_j^{(1)1/2}w_j\|_1\Big], \tag{95}
$$

*where $w = (w_1', \ldots, w_{p_n}')'$ with vectors $w_j = (w_{j1}, \ldots, w_{js_n})' \in \mathbb{R}^{s_n}$. Assume we have the following conditions (a)–(c):*

*(a) $\Sigma_{TT}^{(1)}$ is invertible.*

*(b) $\pi_1\pi_2(1+\lambda_n\|\Lambda_T^{(1)1/2}\beta_T^{(1)}\|_1)(1+\pi_1\pi_2\beta_T^{(1)'}\Sigma_{TT}^{(1)}\beta_T^{(1)})^{-1}(\min_{j\in T}\min_{k\leq s_n}\omega_{jk}^{1/2}|\beta_{jk}|) >$ $\lambda_n\|\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\|_\infty$.*

*(c) $\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\|_\infty \leq 1-\gamma$, for a universal constant $\gamma\in(0,1]$.*

*Denote $\hat{w}$ as $\hat{w} = (\hat{w}_T', \hat{w}_N')' = (\tilde{w}_T', 0')'$ with $\hat{w}_N = 0 \in \mathbb{R}^{(p_n-q_n)s_n}$, and $\hat{w}_T = \tilde{w}_T$ where $\tilde{w}_T$ is defined in Lemma 16. Then we have the following properties:*

*1) $\hat{w}$ is a global minimum of (95).*

*2) sgn($\hat{w}$) = sgn($\beta^{(1)}$).*

*Proof of Lemma 17*: First of all, based on (a), (b) and the definition of $\hat{w}$, it is trivial to deduce that $\text{sgn}(\hat{w}) = \text{sgn}(\beta^{(1)})$, finishing the proof of 2). Moreover, according to the optimization theory, we know that $\hat{w}$ is a global minimum of (95) if and only if

$$\big(\Sigma_{TT}^{(1)} + \pi_1\pi_2\nu_T^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T = \pi_1\pi_2\nu_T^{(1)} - \lambda_n\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}), \tag{96}$$

$$\|\Lambda_N^{(1)-1/2}\big\{\big(\Sigma_{NT}^{(1)} + \pi_1\pi_2\nu_N^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T - \pi_1\pi_2\nu_N^{(1)}\big\}\|_\infty \le \lambda_n, \tag{97}$$

where (96) and (97) serve as the Karush-Kuhn-Tucker conditions. It is apparent that (96) follows from (94). In addition, observe that

$$\begin{aligned}
&\|\Lambda_N^{(1)-1/2}\big\{\big(\Sigma_{NT}^{(1)} + \pi_1\pi_2\nu_N^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T - \pi_1\pi_2\nu_N^{(1)}\big\}\|_\infty \\
=&\|\Lambda_N^{(1)-1/2}\big\{\big(\Sigma_{NT}^{(1)} + \pi_1\pi_2\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T - \pi_1\pi_2\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\big\}\|_\infty \\
=&\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\big\{\big(I + \pi_1\pi_2\beta_T^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T - \pi_1\pi_2\beta_T^{(1)}\big\}\|_\infty,
\end{aligned}$$

where the first and the second equalities follow from (10) in the main paper. For the term $\big(I + \pi_1\pi_2\beta_T^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T$, we have

$$\begin{aligned}
&\big(I + \pi_1\pi_2\beta_T^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T \\
=&\pi_1\pi_2(1 + \lambda_n\|\Lambda_T^{(1)1/2}\beta_T^{(1)}\|_1)\beta_T^{(1)} - \lambda_n\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}) \\
&- \pi_1\pi_2\lambda_n\beta_T^{(1)}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}) \\
=&\pi_1\pi_2\beta_T^{(1)} - \lambda_n\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}),
\end{aligned}$$

where the first equality is by Lemma 16. To this end, based on the above two equations, we deduce that

$$\begin{aligned}
&\|\Lambda_N^{(1)-1/2}\big\{\big(\Sigma_{NT}^{(1)} + \pi_1\pi_2\nu_N^{(1)}\nu_T^{(1)'}\big)\tilde{w}_T - \pi_1\pi_2\nu_N^{(1)}\big\}\|_\infty \\
=&\lambda_n\|\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})\|_\infty \le \lambda_n,
\end{aligned}$$

where the last inequality is based on condition (c). According to the above results, it can be concluded that $\hat{w}$ is a global minimum of (95), which completes the proof. $\square$

**Lemma 18.** *For any $\varrho \in (e^{-n/100}, 1/100)$, define the event $\mathcal{M}_{1n}(\varrho)$ as*

$$\mathcal{M}_{1n}(\varrho) = \Big\{2^{-1}(q_n s_n/n) - 8\{\log(\varrho^{-1})/n\}^{1/2} \leq (\hat{\nu}_T^{(1)\prime} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})/(\hat{\nu}_T^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)}) - 1$$
$$\leq 2(q_n s_n/n) + 16\{\log(\varrho^{-1})/n\}^{1/2}\Big\}.$$

*Assume the condition (a):*

*(a)* $q_n s_n = o(n)$.

*Then we have the following property:*

$$P\{\mathcal{M}_{1n}(\varrho)\} \geq 1 - 2\varrho - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12), \quad \forall \varrho \in (e^{-n/100}, 1/100).$$

*Proof of Lemma 18*: First of all, based on condition (a) and the definition, it is clear to observe that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, we have

$$(n-2) S_{TT}^{(1)} | \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \sim \text{Wishart}(n-2|\Sigma_{TT}^{(1)}), \tag{98}$$

where the degree of freedom of the Wishart distribution is equal to $n-2$. Moreover, it is trivial to verify that conditional on $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, one has the fact that $\hat{\nu}_T^{(1)} | \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$ is independent of $(n-2) S_{TT}^{(1)} | \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$. Together with (98), condition (a) and Theorem 3.2.12 in Muirhead (1982), we reach a conclusion that

$$(n-2)(\hat{\nu}_T^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)})(\hat{\nu}_T^{(1)\prime} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})^{-1} | \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \sim \chi^2_{n - q_n s_n - 1}.$$

Together with (A.2) and (A.3) in Johnstone and Lu (2009), we conclude that for any $t \in [0, 1/2)$,

$$P\big[|(n - q_n s_n - 1)^{-1}(n-2)(\hat{\nu}_T^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)})(\hat{\nu}_T^{(1)\prime} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})^{-1} - 1|$$
$$\geq t | \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \leq 2\exp\{-3(n - q_n s_n - 1)t^2/16\}.$$

For any $\varrho \in (e^{-n/100}, 1/100)$, we plug $t = \{16(n - q_n s_n - 1)^{-1} \log(\varrho^{-1})/3\}^{1/2}$ into the above inequality to obtain

$$P\big[|(n - q_n s_n - 1)^{-1}(n-2)(\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)})(\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})^{-1} - 1| \geq \{16(n - q_n s_n - 1)^{-1} \log(\varrho^{-1})/3\}^{1/2} | \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \leq 2\varrho,$$

which implies that

$$P\big[|(n - q_n s_n - 1)^{-1}(n-2)(\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)})(\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})^{-1} - 1| \leq \{16(n - q_n s_n - 1)^{-1} \log(\varrho^{-1})/3\}^{1/2} | \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \geq 1 - 2\varrho. \tag{99}$$

Therefore, it can be seen that

$$\begin{aligned}
&P\big[|(n - q_n s_n - 1)^{-1}(n-2)(\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)})(\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})^{-1} - 1| \leq \\
&\quad \{16(n - q_n s_n - 1)^{-1} \log(\varrho^{-1})/3\}^{1/2}\big] \\
&\geq \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[|(n - q_n s_n - 1)^{-1}(n-2)(\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)})(\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})^{-1} \\
&\quad - 1| \leq \{16(n - q_n s_n - 1)^{-1} \log(\varrho^{-1})/3\}^{1/2} | \{Y_i = y_i\}_{i=1}^n\big] \cdot P\big[\{Y_i = y_i\}_{i=1}^n\big] \\
&\geq (1 - 2\varrho) \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[\{Y_i = y_i\}_{i=1}^n\big] = (1 - 2\varrho) P(\mathcal{M}_n) \\
&\geq 1 - 2\varrho - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12),
\end{aligned} \tag{100}$$

where the second inequality is by (99), and the last inequality follows from Lemma 3. To this end, based on condition (a), it is straightforward to verify that for any $\varrho \in (e^{-n/100}, 1/100)$,

$$\mathcal{M}_{1n}^*(\varrho) \subseteq \mathcal{M}_{1n}(\varrho), \tag{101}$$

in which $\mathcal{M}_{1n}^*(\varrho) = \big\{|(n - q_n s_n - 1)^{-1}(n-2)(\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\nu}_T^{(1)})(\hat{\nu}_T^{(1)'} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)})^{-1} - 1| \leq \{16(n - q_n s_n - 1)^{-1} \log(\varrho^{-1})/3\}^{1/2}\big\}$. Finally, the assertion follows immediately from (100) and (101). □

**Lemma 19.** *For any $\varrho \in (e^{-n/100}, 1/100)$, define the event $\mathcal{M}_{2n}(\varrho)$ as*

$$\mathcal{M}_{2n}(\varrho) = \Big\{ -(400\pi_1^{-1}\pi_2^{-1})^{1/2}\big[\log(\varrho^{-1})/n + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log(\varrho^{-1})/n\}^{1/2}\big]$$
$$\leq \hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \leq (400\pi_1^{-1}\pi_2^{-1})^{1/2}$$
$$\cdot \big[q_n s_n/n + \log(\varrho^{-1})/n + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log(\varrho^{-1})/n\}^{1/2}\big]\Big\}.$$

*Then we have the following property:*

$$P\{\mathcal{M}_{2n}(\varrho)\} \geq 1 - 2\varrho - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12), \quad \forall \varrho \in (e^{-n/100}, 1/100).$$

*Proof of Lemma 19*: First of all, it is apparent that conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, we have

$$\hat{\nu}_T^{(1)}|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \sim N(\nu_T^{(1)}, nn_1^{-1}n_2^{-1}\Sigma_{TT}^{(1)}),$$

which entails that

$$n_1 n_2 n^{-1}\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \sim \chi^2_{q_n s_n}(n_1 n_2 n^{-1}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}), \tag{102}$$

where $\chi^2_{q_n s_n}(n_1 n_2 n^{-1}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})$ means the noncentral chi-square distribution with $q_n s_n$ degrees of freedom, whose noncentrality parameter has the form $n_1 n_2 n^{-1}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}$. By combining (102) with (8.34) of Lemma 8.1 in Birge (2001), it can be deduced that for any $t > 0$,

$$P\Big[\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \geq (n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) + (n_1^{-1}n_2^{-1}n^2)\cdot$$
$$(2t/n) + 2(n_1^{-1}n_2^{-1}n^2)\{(q_n s_n/n + 2n_1 n_2 n^{-2}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})(t/n)\}^{1/2}$$
$$\Big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big] \leq \exp(-t).$$

By plugging $t = \log(\varrho^{-1})$ into the above inequality, we obtain

$$P\Big[\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \leq (n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) + 2(n_1^{-1}n_2^{-1}n^2)\cdot$$
$$\{\log(\varrho^{-1})/n\} + 2(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n + 2n_1 n_2 n^{-2}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{1/2}\cdot$$
$$\{\log(\varrho^{-1})/n\}^{1/2}\Big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big] \geq 1 - \varrho.$$

Moreover, conditional on $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, we also note that

$$
\begin{aligned}
&(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) + 2(n_1^{-1}n_2^{-1}n^2)\{\log(\varrho^{-1})/n\} \\
&+ 2(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n + 2n_1 n_2 n^{-2}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{1/2}\{\log(\varrho^{-1})/n\}^{1/2} \\
\leq &(400\pi_1^{-1}\pi_2^{-1})^{1/2}\big[q_n s_n/n + \log(\varrho^{-1})/n + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log(\varrho^{-1})/n\}^{1/2}\big],
\end{aligned}
$$

according to the definition of $\mathcal{M}_n$ in Lemma 3. Therefore, based on the above two inequalities, we have

$$
\begin{aligned}
P\Big\{&\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \leq (400\pi_1^{-1}\pi_2^{-1})^{1/2}\big[q_n s_n/n + \log(\varrho^{-1})/n \\
&+ \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log(\varrho^{-1})/n\}^{1/2}\big]\Big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big\} \geq 1 - \varrho. \qquad (103)
\end{aligned}
$$

Analogously, based on (102) and (8.35) of Lemma 8.1 in Birge (2001), it is obvious that for any $t > 0$,

$$
\begin{aligned}
P\Big\{&\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \leq (n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) - 2(n_1^{-1}n_2^{-1}n^2)\cdot \\
&(q_n s_n/n + 2n_1 n_2 n^{-2}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{1/2}(t/n)^{1/2}\Big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big\} \\
&\leq \exp(-t).
\end{aligned}
$$

We then substitute $t = \log(\varrho^{-1})$ into the above inequality to obtain

$$
\begin{aligned}
P\Big[&\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \geq (n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) - 2(n_1^{-1}n_2^{-1}n^2)\cdot \\
&(q_n s_n/n + 2n_1 n_2 n^{-2}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{1/2}\{\log(\varrho^{-1})/n\}^{1/2}\Big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big] \geq 1 - \varrho.
\end{aligned}
$$

Likewise, we note that conditional on $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$,

$$
\begin{aligned}
&(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) - 2(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n + 2n_1 n_2 n^{-2}\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)})^{1/2}\{\log(\varrho^{-1})/n\}^{1/2} \\
\geq &-(400\pi_1^{-1}\pi_2^{-1})^{1/2}\big[\log(\varrho^{-1})/n + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log(\varrho^{-1})/n\}^{1/2}\big].
\end{aligned}
$$

We then derive from the above two inequalities that

$$P\Big\{\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - \nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)} \geq -(400\pi_1^{-1}\pi_2^{-1})^{1/2}\big[\log(\varrho^{-1})/n + \{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log(\varrho^{-1})/n\}^{1/2}\big]\Big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big\} \geq 1-\varrho.$$

Together with (103), we arrive at

$$P\big[\mathcal{M}_{2n}(\varrho)\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \geq 1 - 2\varrho. \tag{104}$$

Finally, we have

$$\begin{aligned} P\{\mathcal{M}_{2n}(\varrho)\} \geq & P\{\mathcal{M}_{2n}(\varrho)\cap\mathcal{M}_n\} = \sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n} P\big[\mathcal{M}_{2n}(\varrho)\big|\{Y_i = y_i\}_{i=1}^n\big]\cdot P\big[\{Y_i = y_i\}_{i=1}^n\big] \\ \geq & (1-2\varrho)\sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n} P\big[\{Y_i = y_i\}_{i=1}^n\big] = (1-2\varrho)P(\mathcal{M}_n) \\ \geq & 1 - 2\varrho - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12), \end{aligned}$$

where the second inequality is by (104), and the last inequality follows from Lemma 3. This finishes the proof. □

**Lemma 20.** *For any $\varrho \in (e^{-n/100}, 1/100)$, define the event $\mathcal{M}_{4n}(\varrho)$ as*

$$\mathcal{M}_{4n}(\varrho) = \bigcap_{j=1}^{q_n s_n}\Big\{\big|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\big| \leq (8\pi_1^{-1}\pi_2^{-1})^{1/2} \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2}\{\log(q_n s_n\varrho^{-1})/n\}^{1/2}\Big\},$$

*where $\{e_j : j \leq q_n s_n\}$ denotes the standard basis for $\mathbb{R}^{q_n s_n}$. Then we have the following property:*

$$P\{\mathcal{M}_{4n}(\varrho)\} \geq 1 - 2\varrho - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12), \quad \forall\varrho\in(e^{-n/100}, 1/100).$$

*Proof of Lemma 20*: First of all, we note that conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap$

$\mathcal{M}_n$

$$\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\big|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\sim N(\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)},n_1^{-1}n_2^{-1}n\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}). \tag{105}$$

Moreover, it can be observed that

$$P\big[\mathcal{M}_{4n}(\varrho)\big|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]\geq\sum_{j=1}^{q_ns_n}P\big[\mathcal{M}_{4nj}(\varrho)\big|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]-(q_ns_n-1),$$

where the events $\mathcal{M}_{4nj}(\varrho)=\Big\{\big|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}-e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\big|\leq(8\pi_1^{-1}\pi_2^{-1})^{1/2}$ $\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2}\{\log(q_ns_n\varrho^{-1})/n\}^{1/2}\Big\}$ for all $j\leq q_ns_n$. Under (105), the concentration inequality entails that for all $j\leq q_ns_n$

$$P\big[\mathcal{M}_{4nj}(\varrho)\big|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]\geq1-2\exp\{-\log(q_ns_n\varrho^{-1})\}=1-2q_n^{-1}s_n^{-1}\varrho.$$

Putting the above two inequalities together leads to

$$P\big[\mathcal{M}_{4n}(\varrho)\big|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]\geq1-2\varrho. \tag{106}$$

Therefore, we have

$$\begin{aligned}&P\{\mathcal{M}_{4n}(\varrho)\}\geq P\{\mathcal{M}_{4n}(\varrho)\cap\mathcal{M}_n\}\\ \geq&(1-2\varrho)\sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n}P\big[\{Y_i=y_i\}_{i=1}^n\big]=(1-2\varrho)P(\mathcal{M}_n)\\ \geq&1-2\varrho-2\exp(-n\pi_1/12)-2\exp(-n\pi_2/12),\end{aligned}$$

where the second inequality is by (106), and the last inequality follows from Lemma 3. This finishes the proof. □

**Lemma 21.** *For any $\varrho\in(e^{-n/100},1/100)$, define the event $\mathcal{M}_{5n}(\varrho)$ as*

$$\begin{aligned}\mathcal{M}_{5n}(\varrho)=&\Big\{\big|\hat{\nu}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})-\nu_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\big|\leq\\ &(8\pi_1^{-1}\pi_2^{-1})^{1/2}\lambda_{\max}^{1/2}(\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\{q_ns_n\log(\varrho^{-1})/n\}^{1/2}\Big\}.\end{aligned}$$

*Then we have the following property:*

$$P\{\mathcal{M}_{5n}(\varrho)\} \geq 1 - 2\varrho - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12), \quad \forall \varrho \in (e^{-n/100}, 1/100).$$

*Proof of Lemma 21*: First of all, we know that conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$

$$\begin{aligned}
&\hat{\nu}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \\
\sim &N\big(\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}), n_1^{-1}n_2^{-1}n\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\big).
\end{aligned}$$

Together with the concentration inequality, we conclude that for any $t > 0$

$$\begin{aligned}
&P\big\{\big|(\hat{\nu}_T^{(1)} - \nu_T^{(1)})'\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\big| \leq t\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big\} \\
\geq &1 - 2\exp\big[-8^{-1}\pi_1\pi_2\{q_n s_n \lambda_{\max}(\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\}^{-1}nt^2\big].
\end{aligned}$$

Plugging $t = (8\pi_1^{-1}\pi_2^{-1})^{1/2}\lambda_{\max}^{1/2}(\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\{q_n s_n \log(\varrho^{-1})/n\}^{1/2}$ into the above inequality yields

$$P\big[\mathcal{M}_{5n}(\varrho)\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \geq 1 - 2\varrho. \tag{107}$$

Finally, we have

$$\begin{aligned}
&P\{\mathcal{M}_{5n}(\varrho)\} \geq P\{\mathcal{M}_{5n}(\varrho) \cap \mathcal{M}_n\} \\
\geq &(1 - 2\varrho) \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\big[\{Y_i = y_i\}_{i=1}^n\big] = (1 - 2\varrho)P(\mathcal{M}_n) \\
\geq &1 - 2\varrho - 2\exp(-n\pi_1/12) - 2\exp(-n\pi_2/12),
\end{aligned}$$

where the second inequality is by (107), and the last inequality follows from Lemma 3. This completes the proof. □

**Lemma 22.** *Assume the following condition (a):*

*(a)* $q_n s_n = o(n)$.

*Then there exists universal constants $c_1 > 0$ and $c_2 > 0$ such that:*

*1)* $P\Big(\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)-1\big|\le$
$c_1\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\Big)\ge 1-c_2\big[\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)\big].$

*2)* $P\Big(\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)-1\big|\le$
$c_1\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\Big)\ge 1-c_2\big[\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)\big].$

*3)* $P\Big(\big|\{sgn(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\}/\{sgn(\beta_T^{(1)})'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\}$
$-1\big|\le c_1\big[q_ns_n/n+\{\log\log(n)/n\}^{1/2}\big]\Big)\ge 1-c_2\big[\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+$
$\exp(-n\pi_2/12)\big].$

*4)* $P\Big(\big|\{sgn(\beta_T^{(1)})'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\}/\{sgn(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})\}$
$-1\big|\le c_1\big[q_ns_n/n+\{\log\log(n)/n\}^{1/2}\big]\Big)\ge 1-c_2\big[\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+$
$\exp(-n\pi_2/12)\big].$

*Recall that* $\{e_j : j\le q_ns_n\}$ *denotes the standard basis for* $\mathbb{R}^{q_ns_n}$*.*

*Proof of Lemma 22*: First of all, according to (98), condition (a) and Theorem 3.2.12 in Muirhead (1982), we know that conditional on any nonempty $\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n$, and for every $j\le q_ns_n$,

$$(n-2)(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\sim\chi^2_{n-q_ns_n-1}.$$

Together with (A.2) and (A.3) in Johnstone and Lu (2009), it can be deduced that for any $t\in[0,1/2)$ and for every $j\le q_ns_n$,

$$P\big[|(n-q_ns_n-1)^{-1}(n-2)(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}$$
$$-1|\ge t|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]\le 2\exp\{-3(n-q_ns_n-1)t^2/16\},$$

which together with condition (a) implies that

$$P\big[|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}-1|$$

$$\leq 4q_ns_n/n+2t|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]\geq 1-2\exp\{-3(n-q_ns_n-1)t^2/16\}$$

$$\geq 1-2\exp(-nt^2/16).$$

Together with the union bound inequality, it can be observed that for any $t\in[0,1/2)$,

$$P\big[\max_{j\leq q_ns_n}|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}-1|$$

$$\leq 4q_ns_n/n+2t|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]\geq 1-2q_ns_n\exp(-nt^2/16).$$

Subsequently, we substitute $t=\{16\log(q_ns_n\log n)/n\}^{1/2}$ into the above inequality to obtain

$$P\big[\max_{j\leq q_ns_n}|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}-1|$$

$$\leq 4q_ns_n/n+8\{\log(q_ns_n\log n)/n\}^{1/2}|\{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\big]$$

$$\geq 1-2\{\log(n)\}^{-1}. \tag{108}$$

It then follows that

$$P\Big(\max_{j\leq q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)-1\big|$$

$$\leq 8\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\Big)$$

$$\geq \sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n}P\Big(\max_{j\leq q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)$$

$$-1\big|\leq 8\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\big|\{Y_i=y_i\}_{i=1}^n\Big)\cdot P\big[\{Y_i=y_i\}_{i=1}^n\big]$$

$$\geq[1-2\{\log(n)\}^{-1}]\sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n}P\big[\{Y_i=y_i\}_{i=1}^n\big]=[1-2\{\log(n)\}^{-1}]P(\mathcal{M}_n)$$

$$\geq 1-2\big[\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)\big],$$

where the second inequality is by (108), and the last inequality follows from Lemma 3. Hence, property 1) is justified by the above inequality. To prove property 2), notice that

under the event $\Big\{\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/$
$(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)-1\big|\le 8\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\Big\}$, it is straightforward to verify that

$$
\begin{aligned}
&\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)-1\big|\\
\le&2\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)-1\big|.
\end{aligned}
$$

Putting the above two inequalities together leads to

$$
\begin{aligned}
&P\Big(\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)-1\big|\le\\
&\quad 16\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\Big)\\
\ge&1-2\big[\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)\big],
\end{aligned}
$$

which completes the proof of property 2). Similar reasoning leads to properties 3) to 4), finishing the proof of the Lemma. □

**Lemma 23.** *Assume the following condition (a):*

*(a)* $q_ns_n=o(n)$.

*Then there exist universal constants* $c_1>0$ *and* $c_2>0$ *such that:*

$$
\begin{aligned}
&P\Big(\bigcap_{j=1}^{q_ns_n}\Big\{|e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}-e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}|\le c_1\big([q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}]\\
&\quad\cdot|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}|+[1+\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}+\{\nu_T^{(1)'}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}\log\log(n)/n\}^{1/2}]^{1/2}\\
&\quad\cdot\{\log(q_ns_n\log n)/n\}^{1/2}\cdot\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2}\big)\Big\}\Big)\\
\ge&1-c_2\big[\{\log(n)\}^{-1}+\exp(-n\pi_1/12)+\exp(-n\pi_2/12)\big].
\end{aligned}
$$

*Proof of Lemma 23*: First of all, we note that for every $j\le q_ns_n$,

$$
|e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}-e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}|\le\Omega_{1j}+\Omega_{2j},\tag{109}
$$

where

$$\Omega_{1j} = |e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}|,$$

$$\Omega_{2j} = |e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)} - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}|.$$

Invoking Lemma 20, it can be deduced that there exist universal constants $c_1 > 0$ and $c_2 > 0$ such that

$$\begin{aligned} &P\Big[\bigcap_{j=1}^{q_n s_n}\Big\{\Omega_{1j} \le c_1\{\log(q_n s_n \log n)/n\}^{1/2}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2}\Big\}\Big] \\ \ge& 1 - c_2\big[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big]. \end{aligned} \tag{110}$$

Regarding the term $\Omega_{2j}$, it can be seen that

$$\Omega_{2j} \le \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}\cdot|\Pi_{1j}|\cdot(1+\Pi_{2j}) + (\Omega_{1j} + |e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}|)\cdot\Pi_{2j},$$

where

$$\begin{aligned} \Pi_{1j} =& \{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\}\{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1} \\ &- \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T^{(1)}\}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1}, \\ \Pi_{2j} =& \big|\{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1} - 1\big|. \end{aligned}$$

For the term $\Pi_{2j}$, it follows from Lemma 22 that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that

$$\begin{aligned} &P\big(\max_{j\le q_n s_n}\Pi_{2j} \le c_3[q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]\big) \\ \ge& 1 - c_4\big[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big]. \end{aligned}$$

To this end, based on the above three inequalities, we conclude that there exist universal constants $c_5 > 0$ and $c_6 > 0$ such that

$$P\Big[\bigcap_{j=1}^{q_n s_n}\Big\{\Omega_{2j} \le c_5\big(|\Pi_{1j}| \cdot \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\} + [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]$$

$$\cdot \{\log(q_n s_n \log n)/n\}^{1/2} \cdot \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2} + [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]$$

$$\cdot |e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\nu_T^{(1)}|\big)\Big\}\Big] \ge 1 - c_6\big[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big]. \tag{111}$$

To bound the term $\Pi_{1j}$, for every $j \le q_n s_n$, we define a $2 \times q_n s_n$ random matrix $\hat{M}_j$ as

$$\hat{M}_j = [\Lambda_T^{(1)1/2}e_j, \hat{\nu}_T]' \in \mathbb{R}^{2\times q_n s_n}.$$

Elementary algebra shows that for every $j \le q_n s_n$,

$$\hat{M}_j S_{TT}^{(1)-1}\hat{M}_j' = \begin{bmatrix} e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j & e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T \\ e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\nu}_T & \hat{\nu}_T'S_{TT}^{(1)-1}\hat{\nu}_T \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

$$\hat{M}_j \Sigma_{TT}^{(1)-1}\hat{M}_j' = \begin{bmatrix} e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j & e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T \\ e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T & \hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{112}$$

Moreover, since $\hat{\nu}_T$ is independent of $S_{TT}^{(1)}$, it can be shown that conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}$, and for every $j \le q_n s_n$,

$$(n-2)(\hat{M}_j S_{TT}^{(1)-1}\hat{M}_j')^{-1}|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\} \sim \text{Wishart}(n - q_n s_n|(\hat{M}_j\Sigma_{TT}^{(1)-1}\hat{M}_j')^{-1}), \tag{113}$$

using Theorem 3.2.11 in Muirhead (1982). To this end, by combining (112), (113) with Theorem 3(d) in Bodnar and Okhrin (2008), it is straightforward to reach a conclusion that for every $j \le q_n s_n$,

$$\{(n - q_n s_n - 3)/\kappa_j\}^{1/2}\Pi_{1j}|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\} \sim t(n - q_n s_n - 3),$$

where $t(n - q_n s_n - 3)$ represents the student t-distribution with $n - q_n s_n - 3$ degrees of freedom, and $\kappa_j = \{\hat{\nu}_T'\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1} - \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\nu}_T\}^2\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-2}$.

Together with Lemma 20 in Kolar and Liu (2015), it is clear that there exist universal constants $c_7 > 0$ and $c_8 > 0$ such that for every $j \le q_n s_n$ and for any $t_j \ge 0$,

$$P\big[|\Pi_{1j}| \ge t_j \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}\big] \le c_7 \exp\{-c_8(n - q_n s_n - 3)\kappa_j^{-1} t_j^2\}$$
$$\le c_7 \exp\big[-2^{-1} c_8 n \{\hat{\nu}_T' \Sigma_{TT}^{(1)-1} \hat{\nu}_T\}^{-1} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\} t_j^2\big],$$

which further implies that

$$P\big[\cap_{j=1}^{q_n s_n} \{|\Pi_{1j}| \le t_j\} \big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}\big]$$
$$\ge 1 - \sum_{j=1}^{q_n s_n} c_7 \exp\big[-2^{-1} c_8 n \{\hat{\nu}_T' \Sigma_{TT}^{(1)-1} \hat{\nu}_T\}^{-1} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\} t_j^2\big].$$

By plugging $t_j = c_9 \{\hat{\nu}_T' \Sigma_{TT}^{(1)-1} \hat{\nu}_T\}^{1/2} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\}^{-1/2} \{\log(q_n s_n \log n)/n\}^{1/2}$ with $c_9 = (2c_8^{-1})^{1/2}$ into the above inequality, it can be obtained that

$$P\Big[\bigcap_{j=1}^{q_n s_n} \Big\{|\Pi_{1j}| \le c_9 \{\hat{\nu}_T' \Sigma_{TT}^{(1)-1} \hat{\nu}_T\}^{1/2} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\}^{-1/2} \{\log(q_n s_n \log n)/n\}^{1/2}$$
$$\Big\} \Big| \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}\Big] \ge 1 - c_7 \{\log(n)\}^{-1}. \tag{114}$$

It then follows that

$$P\Big[\bigcap_{j=1}^{q_n s_n} \Big\{|\Pi_{1j}| \le c_9 \{\hat{\nu}_T' \Sigma_{TT}^{(1)-1} \hat{\nu}_T\}^{1/2} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\}^{-1/2} \{\log(q_n s_n \log n)/n\}^{1/2}\Big\}\Big]$$
$$\ge \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} \sum_{\hat{\nu}_T \in \mathcal{M}_n} P\Big[\bigcap_{j=1}^{q_n s_n} \Big\{|\Pi_{1j}| \le c_9 \{\hat{\nu}_T' \Sigma_{TT}^{(1)-1} \hat{\nu}_T\}^{1/2} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\}^{-1/2}$$
$$\{\log(q_n s_n \log n)/n\}^{1/2}\Big\} \Big| \{Y_i = y_i\}_{i=1}^n \cap \{\hat{\nu}_T\}\Big] \cdot P\big[\{Y_i = y_i\}_{i=1}^n \cap \{\hat{\nu}_T\}\big]$$
$$\ge [1 - c_7 \{\log(n)\}^{-1}] \cdot \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} \sum_{\hat{\nu}_T \in \mathcal{M}_n} P\big[\{Y_i = y_i\}_{i=1}^n \cap \{\hat{\nu}_T\}\big] = [1 - c_7 \{\log(n)\}^{-1}] \cdot P(\mathcal{M}_n)$$
$$\ge 1 - c_{10}\big[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\big],$$

for some universal constant $c_{10} > 0$, where the second inequality is by (114). Together with

Lemma 19, it is seen that there exist universal constants $c_{11} > 0$ and $c_{12} > 0$ such that,

$$\begin{aligned} P\Big[ \bigcap_{j=1}^{q_n s_n} \Big\{ &|\Pi_{1j}| \le c_{11} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\}^{-1/2} \{\log(q_n s_n \log n)/n\}^{1/2} \\ &\cdot \big( q_n s_n/n + \log\log(n)/n + [1 + q_n s_n/n + \{\log\log(n)/n\}^{1/2}] \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\} \\ &+ \{\log\log(n)/n\}^{1/2} \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}\}^{1/2} \big)^{1/2} \Big\} \Big] \\ \ge & 1 - c_{12} \big[ \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12) \big]. \end{aligned}$$

Together with (111), it is clear that there exist universal constants $c_{13} > 0$ and $c_{14} > 0$ such that,

$$\begin{aligned} P\Big( \bigcap_{j=1}^{q_n s_n} \Big\{ &\Omega_{2j} \le c_{13} \big( [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}] \cdot |e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| \\ &+ [q_n s_n/n + \log\log(n)/n + \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} + \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2} \\ &\cdot \{\log(q_n s_n \log n)/n\}^{1/2} \cdot \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\}^{1/2} \big) \Big\} \Big) \\ \ge & 1 - c_{14} \big[ \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12) \big]. \end{aligned}$$

Together with (109) and (110), it is not difficult to verify that there exist universal constants $c_{15} > 0$ and $c_{16} > 0$ such that,

$$\begin{aligned} P\Big( \bigcap_{j=1}^{q_n s_n} \Big\{ &|e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\nu}_T^{(1)} - e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| \le c_{15} \big( [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}] \\ &\cdot |e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \nu_T^{(1)}| + [1 + \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} + \{\nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \nu_T^{(1)} \log\log(n)/n\}^{1/2}]^{1/2} \\ &\cdot \{\log(q_n s_n \log n)/n\}^{1/2} \cdot \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\}^{1/2} \big) \Big\} \Big) \\ \ge & 1 - c_{16} \big[ \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12) \big], \end{aligned}$$

which completes the proof. □

**Lemma 24.** *Assume the following conditions (a)–(b):*

*(a)* $q_n^2 s_n^2 \log(q_n s_n) = o(n)$.

*(b)* $c_1 \le \lambda_{\min}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \le \lambda_{\max}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \le c_2$, *for some universal constants* $0 < c_1 < c_2$.

*Then there exist universal constants* $c_3 > 0$ *and* $c_4 > 0$ *such that:*

$$
\begin{aligned}
P\Big[\bigcap_{j=1}^{q_n s_n}\Big\{&|e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}sgn(\beta_T^{(1)}) - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})| \\
&\le c_3\big[\{q_n s_n\log(q_n s_n)/n\}^{1/2} + \{q_n s_n\log\log(n)/n\}^{1/2}\big] \\
&+ c_3|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})|\cdot\big[q_n s_n/n + \{\log(q_n s_n\log n)/n\}^{1/2}\big]\Big\}\Big] \\
\ge& 1 - c_4[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].
\end{aligned}
$$

*Proof of Lemma 24*: First of all, we note that for every $j \le q_n s_n$,

$$
|e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}) - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})| \le \Omega_{1j} + \Omega_{2j}, \tag{115}
$$

where

$$
\begin{aligned}
\Omega_{1j} &= |e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}) - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})|, \\
\Omega_{2j} &= |e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)}) - e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})|.
\end{aligned}
$$

For the term $\Omega_{1j}$, it is apparent to see that for every $j \le q_n s_n$,

$$
\Omega_{1j} \le c_1^{-1}(1+\Pi_{1j})\cdot|\Pi_{2j}| + |e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})|\cdot\Pi_{1j}, \tag{116}
$$

where $c_1$ is defined in condition (b), and

$$
\begin{aligned}
\Pi_{1j} =& |\{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1} - 1|, \\
\Pi_{2j} =& \{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})\}\{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1} \\
&- \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}(\beta_T^{(1)})\}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1}.
\end{aligned}
$$

To bound the term $\Pi_{1j}$, invoking Lemma 22, it can be seen that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that with probability at least $1 - c_3[\{\log(n)\}^{-1} +$

$\exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\max_{j \le q_n s_n} \Pi_{1j} \le c_4[q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]. \tag{117}$$

To bound the term $\Pi_{2j}$, based on similar argument as in the proof of Lemma 23, it can be shown that there exist universal constants $c_5 > 0$ and $c_6 > 0$ such that conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, and for any $t \ge 0$,

$$P\big[\cap_{j=1}^{q_n s_n} \{|\Pi_{2j}| \le t\}\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \ge 1 - c_5 q_n s_n \exp\{-c_6 n (q_n s_n)^{-1} t^2\}.$$

By setting $c_7 = c_6^{-1/2}$ and plugging $t = c_7\{q_n s_n \log(q_n s_n \log n)/n\}^{1/2}$ into the above inequality, it can be obtained that

$$P\big[\max_{j \le q_n s_n} |\Pi_{2j}| \le c_7\{q_n s_n \log(q_n s_n \log n)/n\}^{1/2}\big|\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\big] \ge 1 - c_5\{\log(n)\}^{-1}.$$

Together with Lemma 3, there exist universal constants $c_8 > 0$ and $c_9 > 0$ such that with probability at least $1 - c_8[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

$$\max_{j \le q_n s_n} \Pi_{2j} \le c_9\{q_n s_n \log(q_n s_n \log n)/n\}^{1/2}. \tag{118}$$

By combining (117), (118) with (116), it is seen that there exist universal constants $c_{10} > 0$ and $c_{11} > 0$ such that

$$\begin{aligned} P\Big[\bigcap_{j=1}^{q_n s_n} \Big\{&\Omega_{1j} \le c_{10}\{q_n s_n \log(q_n s_n \log n)/n\}^{1/2} + c_{10}|e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| \\ &\cdot [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]\Big\}\Big] \\ \ge& 1 - c_{11}[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]. \end{aligned} \tag{119}$$

To bound the term $\Omega_{2j}$, it can be verified that

$$\max_{j \le q_n s_n} \Omega_{2j} \le (q_n s_n)^{1/2} \|\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} - I_{q_n s_n}\|_{\max} \cdot \|\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}\|_2.$$

Together with Lemma 5 and Lemma 8, it is seen that there exist universal constants $c_{12}, c_{13} > 0$ such that

$$
\begin{aligned}
&P\big[\max_{j\le q_n s_n} \Omega_{2j} \le c_{12}\{q_n s_n \log(q_n s_n)/n\}^{1/2}\big] \\
\ge &1 - c_{13}[(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)].
\end{aligned}
$$

Together with (115) and (119), the assertion holds trivially, which completes the proof. $\square$

# References


Birge, L. (2001). An alternative point of view on lepski's method. *State of the Art in Probability and Statistics*, 36:113–133.

Bodnar, T. and Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned wishart distributions. *Journal of Multivariate Analysis*, 99(10):2389–2405.

Johnstone, I. M. and Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association*, 104(486):682–693.

Kolar, M. and Liu, H. (2015). Optimal feature selection in high-dimensional discriminant analysis. *IEEE Transactions on Information Theory*, 61(2):1063–1083.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley Series in Probability and Statistics. Wiley.

Ning, Y. and Liu, H. (2017). A general theory of hypothesis tests and confidence regions for sparse high dimensional models. *The Annals of Statistics*, 45(1):158–195.

Shao, J., Wang, Y., Deng, X., and Wang, S. (2011). Sparse linear discriminant analysis by thresholding for high dimensional data. *The Annals of Statistics*, 39(2):1241–1265.