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 $j \quad p_n$, 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

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

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

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

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

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

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

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

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

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

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

Similar to the constructions of the vectors $^{(1)}_{T}$, $^{(1)}_{1,T}$, $^{(1)}_{2,T}$, and $^{(1)}_{T}$, we define $^{(1)}_{i,T}$, $\hat{}^{(1)}_{1,T}$, $\hat{}^{(1)}_{2,T}$, and $\hat{}^{(1)}_{T}$ as restricting the vectors $^{(1)}_{i}$, $\hat{}^{(1)}_{1}$, $\hat{}^{(1)}_{2}$, and $\hat{}^{(1)}$ to the discriminant set T. Given index sets T and N, we define several sample covariance matrices as

$$S^{(1)} = f(n_1 \quad 1)S^{(1)}_1 + (n_2 \quad 1)S^{(1)}_2 g=(n \quad 2);$$

$$S^{(1)}_{TT} = f(n_1 \quad 1)S^{(1)}_{1,TT} + (n_2 \quad 1)S^{(1)}_{2,TT} g=(n \quad 2);$$

$$S^{(1)}_{NT} = f(n_1 \quad 1)S^{(1)}_{1,NT} + (n_2 \quad 1)S^{(1)}_{2,NT} g=(n \quad 2);$$

where

$$S^{(1)}_1 = \sum_{i\in H_1} (^{(1)}_{i} \quad \hat{}^{(1)}_{1})(^{(1)}_{i} \quad \hat{}^{(1)}_{1})'=(n_1 \quad 1);$$

$$S^{(1)}_2 = \sum_{i\in H_2} (^{(1)}_{i} \quad \hat{}^{(1)}_{2})(^{(1)}_{i} \quad \hat{}^{(1)}_{2})'=(n_2 \quad 1);$$

$$S^{(1)}_{1,TT} = \sum_{i\in H_1} (^{(1)}_{i,T} \quad \hat{}^{(1)}_{1,T})(^{(1)}_{i,T} \quad \hat{}^{(1)}_{1,T})'=(n_1 \quad 1);$$

$$S^{(1)}_{2,TT} = \sum_{i\in H_2} (^{(1)}_{i,T} \quad \hat{}^{(1)}_{2,T})(^{(1)}_{i,T} \quad \hat{}^{(1)}_{2,T})'=(n_2 \quad 1);$$

$$S^{(1)}_{1,NT} = \sum_{i\in H_1} (^{(1)}_{i,N} \quad \hat{}^{(1)}_{1,N})(^{(1)}_{i,T} \quad \hat{}^{(1)}_{1,T})'=(n_1 \quad 1);$$

$$S^{(1)}_{2,NT} = \sum_{i\in H_2} (^{(1)}_{i,N} \quad \hat{}^{(1)}_{2,N})(^{(1)}_{i,T} \quad \hat{}^{(1)}_{2,T})'=(n_2 \quad 1):$$

Similar to the definitions of $^{(1)}_{1}$, $^{(1)}_{2}$, $^{(1)}$, $^{(1)}_{1,T}$, $^{(1)}_{2,T}$, and $^{(1)}_{T}$, we denote for any ` $= 1; 2$,

$$^{(2)}_{\ell} = E(^{(2)} jY = `) = (\tilde{}^{(2)\prime}_{\ell 1}; \dots; \tilde{}^{(2)\prime}_{\ell p_n})';$$

$$\tilde{}^{(2)}_{\ell j} = E(\tilde{}^{(2)}_{j} jY = `) = (_{\ell j, s_n+1}; _{\ell j, s_n+2}; \dots)' 2 \mathbb{R}^{\infty}; \quad j = 1; \dots; p_n;$$

$^{(2)}_{\ell,T}$: formed by stacking $f \tilde{}^{(2)}_{\ell j} : j \; 2 \; T g$ in a column,

$$^{(2)} = \; ^{(2)}_{2} \quad ^{(2)}_{1}; \quad ^{(2)}_{T} = \; ^{(2)}_{2,T} \quad ^{(2)}_{1,T}.$$

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

$$^{*(1)} = ({}_1^{*(1)'}, \ldots, {}_{p_n}^{*(1)'})' \text{ with each } {}_j^{*(1)} = ({}_{j1}^{*}, \ldots, {}_{js_n}^{*})',$$

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

$$^{*(2)} = ({}_1^{*(2)'}, \ldots, {}_{p_n}^{*(2)'})' \text{ with each } {}_j^{*(2)} = ({}_{j,s_n+1}^{*}, {}_{j,s_n+2}^{*}, \ldots)',$$

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

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

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

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

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

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

$$R(^*) = R^\circ(^{(1)}) = ({}_1 + {}_2\Omega_1)({}_1 + {}_2\Omega_2)^{-1}\Omega_3, \tag{3}$$

where

$$\Omega_1 = \Phi(\ \ \Delta{=}2 + \log(\ _1{=}\ _2){=}\Delta){=}\Phi(\ \ \Delta{=}2 + \log(\ _2{=}\ _1){=}\Delta);$$

$$\Omega_2 = \Phi(\ \ \Delta^{(1)}{=}2 + \log(\ _1{=}\ _2){=}\Delta^{(1)}){=}\Phi(\ \ \Delta^{(1)}{=}2 + \log(\ _2{=}\ _1){=}\Delta^{(1)});$$

$$\Omega_3 = \Phi(\ \ \Delta{=}2 + \log(\ _2{=}\ _1){=}\Delta){=}\Phi(\ \ \Delta^{(1)}{=}2 + \log(\ _2{=}\ _1){=}\Delta^{(1)}):$$

For the term Ω_1, it can be rewritten as

$$\Omega_1 = \Phi(\ \ \%_n^{1/2}(1 + \#_n)){=}\Phi(\ \ \%_n^{1/2}); \tag{4}$$

where $\%_n = f\Delta{=}2 \quad \log(\ _2{=}\ _1){=}\Delta g^2$ and $\#_n = 4\log(\ _2{=}\ _1){=}f\Delta^2 \quad 2\log(\ _2{=}\ _1)g$. Since $\%_n\ !\ 1$ and $\%_n\#_n\ !\ \log(\ _2{=}\ _1)$ under (1), we immediately conclude that

$$\Omega_1\ !\ \ _2{=}\ _1; \tag{5}$$

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

$$\Omega_2\ !\ \ _2{=}\ _1: \tag{6}$$

For the term Ω_3, it can be expressed as

$$\Omega_3 = \Phi(\ \ \%_n^{1/2}(1 + \tilde{\#}_n)){=}\Phi(\ \ \%_n^{1/2}); \tag{7}$$

where $\%_n = f\Delta^{(1)}{=}2 \quad \log(\ _2{=}\ _1){=}\Delta^{(1)}g^2$ and $\tilde{\#}_n = [f\Delta\Delta^{(1)}{+}2\log(\ _2{=}\ _1)g(\Delta \quad \Delta^{(1)})]{=}f\Delta\Delta^{(1)2}$ $2\log(\ _2{=}\ _1)\Delta g$. Based on (2) and (A3), one can show that

$$\%_n\ !\ 1; \quad \%_n\tilde{\#}_n\ !\ 0:$$

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

$$\Omega_3\ !\ 1:$$

Together with (3), (5) and (6), we have $R(\ ^*){=}R^\circ(\ ^{(1)})\ !\ 1$, 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

$$jY = 1 \quad N(\ _1; \Sigma); \qquad jY = 2 \quad N(\ _2; \Sigma);$$

which implies

$$\Sigma^{\dagger 1/2} \ jY = 1 \quad N(\Sigma^{\dagger 1/2} \ _1; I); \qquad \Sigma^{\dagger 1/2} \ jY = 2 \quad N(\Sigma^{\dagger 1/2} \ _2; I):$$

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

$$f_z(zjY = i) \ / \ \exp f \ \ 2^{-1}(z \quad \Sigma^{\dagger 1/2} \ _i)'(z \quad \Sigma^{\dagger 1/2} \ _i)g; \quad \text{for} \ \ i = 1; 2:$$

By change of variables, the conditional density functions of $\ = \Sigma^2$

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

$$\tilde{v}_T = fn_1n_2n^{-1}(n \quad 2)^{-1}gf1 + \ _n \hat{\ }^{(1)'}_T S^{(1)-1}_{TT}\hat{\Lambda}^{(1)1/2}_T\mathrm{sgn}(\ ^{(1)}_T)g \ 1+$$

$$fn_1n_2n^{-1}(n \quad 2)^{-1}g\hat{\ }^{(1)'}_T S^{(1)-1}_{TT}\hat{\ }^{(1)}_T \ ^{-1}S^{(1)-1}_{TT}\hat{\ }^{(1)}_T \quad _nS^{(1)-1}_{TT}\hat{\Lambda}^{(1)1/2}_T\mathrm{sgn}(\ ^{(1)}_T):$$

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

$$P \ \ fn_1n_2n^{-1}(n \quad 2)^{-1}g\hat{\ }^{(1)}_T \quad fS^{(1)}_{TT} + n_1n_2n^{-1}(n \quad 2)^{-1}\hat{\ }^{(1)}_T\hat{\ }^{(1)'}_T g\tilde{v}_T =$$

$$_n\hat{\Lambda}^{(1)1/2}_T\mathrm{sgn}(\tilde{v}_T) \quad 1 \quad c_5[f(p_n \quad q_n)s_ng^{-1} + (q_ns_n)^{-1} + f\log(n)g^{-1}+$$

$$\exp(\ n \ \ _1{=}12) + \exp(\ n \ \ _2{=}12)]; \tag{10}$$

and

$$P \quad \hat{\Lambda}^{(1)-1/2}_N \ fn_1n_2n^{-1}(n \quad 2)^{-1}g\hat{\ }^{(1)}_N \quad fS^{(1)}_{NT} + n_1n_2n^{-1}(n \quad 2)^{-1}\hat{\ }^{(1)}_N\hat{\ }^{(1)'}_T g$$

$$\tilde{v}_T \ \ _\infty \quad _n \quad 1 \quad c_6[f(p_n \quad q_n)s_ng^{-1} + (q_ns_n)^{-1} + f\log(n)g^{-1}+$$

$$\exp(\ n \ \ _1{=}12) + \exp(\ n \ \ _2{=}12)]: \tag{11}$$

Note that the random quantity $S^{(1)}_{NT}$ can be expressed as $S^{(1)}_{NT} = f(n_1 \quad 1)S^{(1)}_{1,NT} + (n_2 \quad 1)S^{(1)}_{2,NT}g{=}(n \quad 2)$, where $S^{(1)}_{1,NT} = \sum_{i\in H_1}(\ ^{(1)}_{i,N} \quad \hat{\ }^{(1)}_{1,N})(\ ^{(1)}_{i,T} \quad \hat{\ }^{(1)}_{1,T})'{=}(n_1 \quad 1)$ and $S^{(1)}_{2,NT} = \sum_{i\in H_2}(\ ^{(1)}_{i,N} \quad \hat{\ }^{(1)}_{2,N})(\ ^{(1)}_{i,T} \quad \hat{\ }^{(1)}_{2,T})'{=}(n_2 \quad 1)$. Since $\tilde{v}_T$ 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

$$\hat{\Lambda}^{(1)-1/2}_N \ fn_1n_2n^{-1}(n \quad 2)^{-1}g\hat{\ }^{(1)}_N \quad fS^{(1)}_{NT} + n_1n_2n^{-1}(n \quad 2)^{-1} \tag{12}$$

$$\hat{\ }^{(1)}_N\hat{\ }^{(1)'}_T g\tilde{v}_T \ \ _\infty \quad 1 + k\hat{\Lambda}^{(1)-1/2}_N\Lambda^{(1)1/2}_N \quad I_{(p_n-q_n)s_n}k_{\max} \quad k\Psi k_\infty;$$

where $\Psi = \Lambda^{(1)-1/2}_N \ fS^{(1)}_{NT} + n_1n_2n^{-1}(n \quad 2)^{-1}\hat{\ }^{(1)}_N\hat{\ }^{(1)'}_T g\tilde{v}_T \quad fn_1n_2n^{-1}(n \quad 2)^{-1}g\hat{\ }^{(1)}_N$. By definition, conditional on any nonempty set $\{Y_i = y_i\}_{i=1}^n \cap M_n$,

$$
\begin{aligned}
&(n-2)\Lambda^{(1)-1/2}S^{(1)}\Lambda^{(1)-1/2} \mid \{Y_i = y_i\}_{i=1}^n \cap M_n \\
&\sim \text{Wishart}(n-2 \mid \Lambda^{(1)-1/2}\Sigma^{(1)}\Lambda^{(1)-1/2}). \qquad (13)
\end{aligned}
$$

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

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

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.

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

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

$$
\{Z_l\}_{l=1}^{n-2} \perp \hat{\mu}^{(1)}. \qquad (15)
$$

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

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

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

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

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

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

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

$$\begin{aligned} Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2} Z_T &= [WZ_1, \ldots, WZ_{n-2}], \\ Z_T &= [W^* Z_1, \ldots, W^* Z_{n-2}], \end{aligned} \tag{20}$$

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

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

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

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

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

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

$$
\begin{aligned}
\hat{}_T^{(1)} \quad &? \quad \{Z_T; Z_N \quad \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T\};\\
\hat{}_T^{(1)} \quad &? \quad Z_N \quad \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T; \qquad (23)\\
\hat{}_T^{(1)} \quad &? \quad Z_T.
\end{aligned}
$$

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

$$
\{\hat{}_T^{(1)}; Z_T\} \quad ? \quad Z_N \quad \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T. \qquad (24)
$$

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

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

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

$$
\begin{bmatrix}
\Lambda_N^{(1)-1/2}(\hat{}_N^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{}_T^{(1)}) \\
\Lambda_T^{(1)-1/2}\hat{}_T^{(1)}
\end{bmatrix}
\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \quad N
\begin{bmatrix}
\Lambda_N^{(1)-1/2}({}_N^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1} {}_T^{(1)}) \\
\Lambda_T^{(1)-1/2} {}_T^{(1)}
\end{bmatrix};
\qquad (26)
$$

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

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

$$
\hat{}_T^{(1)} \quad ? \quad \hat{}_N^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\hat{}_T^{(1)}. \qquad (27)
$$

Based on (18), it is seen that conditional on any nonempty set $fY_i = y_i g_{i=1}^{n} \setminus M_n$,

$$
\begin{aligned}
Z_T &\;?\; f\,\hat{\ }^{(1)}_{T};\ \hat{\ }^{(1)}_{N} \quad \Sigma^{(1)}_{NT}\Sigma^{(1)-1}_{TT}\,\hat{\ }^{(1)}_{T} g; \\
Z_T &\;?\; \hat{\ }^{(1)}_{N} \quad \Sigma^{(1)}_{NT}\Sigma^{(1)-1}_{TT}\,\hat{\ }^{(1)}_{T}; \qquad (28)\\
Z_T &\;?\; \hat{\ }^{(1)}_{T}:
\end{aligned}
$$

Together with (27) yields that conditional on any nonempty set $fY_i = y_i g_{i=1}^{n} \setminus M_n$,

$$
f\,\hat{\ }^{(1)}_{T}; Z_T g \;?\; \hat{\ }^{(1)}_{N} \quad \Sigma^{(1)}_{NT}\Sigma^{(1)-1}_{TT}\,\hat{\ }^{(1)}_{T}: \qquad (29)
$$

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

$$
\Psi = \Pi_1 \quad \Pi_2 \quad \Pi_3 \quad \Pi_4; \qquad (30)
$$

with

$$
\begin{aligned}
\Pi_1 &= (n \quad 2)^{-1}(Z_N \quad \Lambda^{(1)-1/2}_{N}\Sigma^{(1)}_{NT}\Sigma^{(1)-1}_{TT}\Lambda^{(1)1/2}_{T} Z_T) Z'_T \Lambda^{(1)1/2}_{T}\tilde{v}_T; \\
\Pi_2 &= \hat{\#}\Lambda^{(1)-1/2}_{N}(\hat{\ }^{(1)}_{N} \quad \Sigma^{(1)}_{NT}\Sigma^{(1)-1}_{TT}\,\hat{\ }^{(1)}_{T}); \\
\Pi_3 &= \quad {}_n\Lambda^{(1)-1/2}_{N}\Sigma^{(1)}_{NT}\Sigma^{(1)-1}_{TT}\Lambda^{(1)1/2}_{T}(\Lambda^{(1)-1/2}_{T}\hat{\Lambda}^{(1)1/2}_{T} \quad I_{q_n s_n})\mathrm{sgn}(\ {}^{(1)}_{T}); \\
\Pi_4 &= \quad {}_n\Lambda^{(1)-1/2}_{N}\Sigma^{(1)}_{NT}\Sigma^{(1)-1}_{TT}\Lambda^{(1)1/2}_{T}\mathrm{sgn}(\ {}^{(1)}_{T});
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\#} = & f n_1 n_2 n^{-1}(n \quad 2)^{-1} g f 1 + \ {}_n\,\hat{\ }^{(1)'}_{T} S^{(1)-1}_{TT}\hat{\Lambda}^{(1)1/2}_{T}\mathrm{sgn}(\ {}^{(1)}_{T}) g \\
& 1 + f n_1 n_2 n^{-1}(n \quad 2)^{-1} g\,\hat{\ }^{(1)'}_{T} S^{(1)-1}_{TT}\,\hat{\ }^{(1)}_{T} \quad {}^{-1}:
\end{aligned}
$$

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 $1 \quad c_7[(q_n s_n)^{-1} + f\log(n) g^{-1} + \exp(\ \ n \ \ {}_1=12) + \exp(\ \ n \ \ {}_2=12)]$,

$$
c_8 f \ {}^{(1)'}_{T}\Sigma^{(1)-1}_{TT}\ {}^{(1)}_{T} g^{-1} \quad \hat{\#} \quad c_9 f \ {}^{(1)'}_{T}\Sigma^{(1)-1}_{TT}\ {}^{(1)}_{T} g^{-1}: \qquad (31)
$$

For the term Π_1, it can be decomposed into

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

where

$$\Upsilon_1 = \hat{\omega}(n-2)^{-1}(Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T)Z_T'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\beta}_T^{(1)},$$
$$\Upsilon_2 = \lambda_n(n-2)^{-1}(Z_N - \Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}Z_T)Z_T'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}).$$

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 $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\beta}_T^{(1)}, Z_T\}$ and for any $j \le (p_n - q_n)s_n$,

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

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

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

with each

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

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

$$P\left(\|\Upsilon_1\|_\infty \ge t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\beta}_T^{(1)}, Z_T\}\right)$$
$$\le 2(p_n - q_n)s_n \exp\left(-4^{-1}\hat{\omega}^{-2}\{\hat{\beta}_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\beta}_T^{(1)}\}^{-1}nt^2\right).$$

with each

$$\Xi_j = {}_n^2(n \quad 2)^{-1}\{e_j'\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}e_j\}\{\mathrm{sgn}({}_T^{(1)})'$$
$$\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}({}_T^{(1)})\}$$
$${}_n^2(n \quad 2)^{-1}\{\mathrm{sgn}({}_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}({}_T^{(1)})\}.$$

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

$$P \ \|\Upsilon_2\|_\infty \quad t \ \{Y_i = y_i\}_{i=1}^n \cap M_n \cap \{\hat{}_T^{(1)}, Z_T\}$$
$$2(p_n \quad q_n)s_n \exp \quad 4^{-1} \ {}_n^{-2}\{\mathrm{sgn}({}_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}({}_T^{(1)})\}^{-1}nt^2 \ .$$

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

$$P \ \|\Upsilon_2\|_\infty \quad [8 \ {}_n^2\{\mathrm{sgn}({}_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}({}_T^{(1)})\}]^{1/2}$$
$$[\log\{(p_n \quad q_n)s_n\}/n]^{1/2} \ \{Y_i = y_i\}_{i=1}^n \cap M_n \cap \{\hat{}_T^{(1)}, Z_T\}$$
$$1 \quad 2\{(p_n \quad q_n)s_n\}^{-1}.$$

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

$$P \ \|\Upsilon_2\|_\infty \quad [8 \ {}_n^2\{\mathrm{sgn}({}_T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}({}_T^{(1)})\}]^{1/2}$$
$$[\log\{(p_n \quad q_n)s_n\}/n]^{1/2}$$
$$1 \quad c_{13}[\{(p_n \quad q_n)s_n\}^{-1} + \exp(\ n \ {}_1/12) + \exp(\ n \ {}_2/12)],$$

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 $1 \quad c_{14}[\{(p_n \quad q_n)s_n\}^{-1} + (q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(\ n \ {}_1/12) + \exp(\ n \ {}_2/12)]$,

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

For the term Π_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 $1 \quad c_{21} f(q_n s_n)^{-1} + \exp(\quad n \quad {}_1{=}12) + \exp(\quad n \quad {}_2{=}12) g$, we have $k\Pi_3 k_\infty \quad c_{22} f q_n s_n \log(q_n s_n){=}n g^{1/2} \quad {}_n$. Together with (37), (36) and (30), there exists a universal constant $c_{23} > 0$ such that with probability at least $1 \quad c_{23}[f(p_n \quad q_n) s_n g^{-1} + (q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(\quad n \quad {}_1{=}12) + \exp(\quad n \quad {}_2{=}12)]$, we have $k\Psi k_\infty \quad (1 \quad {=}2) \quad {}_n$. 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 $\tilde{v} = (\tilde{v}_T'; 0')' \; 2 \; \mathbb{R}^{p_n s_n}$, 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 $1 \quad c_{24} f(q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(\quad n \quad {}_1{=}12) + \exp(\quad n \quad {}_2{=}12) g$,

$$R^{\diamond}(\hat{v}) = R^{\diamond}(\tilde{v}) = \quad {}_1\Omega_1 + \quad {}_2\Omega_2; \tag{38}$$

where

$$\Omega_1 = \Phi \qquad \tilde{v}_T'(\hat{\ }^{(1)}_{1,T} \qquad {}^{(1)}_{1,T}) \quad 2^{-1}\tilde{v}_T' \hat{\ }^{(1)}_T + f\tilde{v}_T' S^{(1)}_{TT}\tilde{v}_T g f \tilde{v}_T' \hat{\ }^{(1)}_T g^{-1} f \log(n_2{=}n_1) g$$
$$f\tilde{v}_T'\Sigma^{(1)}_{TT}\tilde{v}_T g^{-1/2} \quad ;$$
$$\Omega_2 = \Phi \qquad \tilde{v}_T'({}^{(1)}_{2,T} \qquad \hat{\ }^{(1)}_{2,T}) \quad 2^{-1}\tilde{v}_T' \hat{\ }^{(1)}_T \qquad f\tilde{v}_T' S^{(1)}_{TT}\tilde{v}_T g f \tilde{v}_T' \hat{\ }^{(1)}_T g^{-1} f \log(n_2{=}n_1) g$$
$$f\tilde{v}_T'\Sigma^{(1)}_{TT}\tilde{v}_T g^{-1/2} \quad :$$

Also recalling from (11) of the main paper that

$$R^{\circ}(\quad {}^{(1)}) = \quad {}_1\Omega_1^* + \quad {}_2\Omega_2^*; \tag{39}$$

with

$$\Omega_1^* = \Phi \qquad 2^{-1} f \quad {}^{(1)'}_T \Sigma^{(1)-1}_{TT} \quad {}^{(1)}_T g^{1/2} + \log(\quad {}_2{=} \quad {}_1) f \quad {}^{(1)'}_T \Sigma^{(1)-1}_{TT} \quad {}^{(1)}_T g^{-1/2} \quad ;$$
$$\Omega_2^* = \Phi \qquad 2^{-1} f \quad {}^{(1)'}_T \Sigma^{(1)-1}_{TT} \quad {}^{(1)}_T g^{1/2} \qquad \log(\quad {}_2{=} \quad {}_1) f \quad {}^{(1)'}_T \Sigma^{(1)-1}_{TT} \quad {}^{(1)}_T g^{-1/2} \quad :$$

We denote a_n, b_n, X_n and U_n as

$$a_n = 4^{-1} \ {}_T^{(1)\prime} \Sigma_{TT}^{(1)-1} \ {}_T^{(1)}; \qquad b_n = \log(\ _2 = \ _1) f \ {}_T^{(1)\prime} \Sigma_{TT}^{(1)-1} \ {}_T^{(1)} g^{-1/2};$$

$$X_n = f 2\tilde{v}_T'(\hat{}^{(1)}_{1,T} \qquad {}^{(1)}_{1,T}) + \tilde{v}_T' \hat{}^{(1)}_T g f \tilde{v}_T' \Sigma^{(1)}_{TT} \tilde{v}_T g^{-1/2} f \ {}^{(1)\prime}_T \Sigma^{(1)-1}_{T1/} g^{-1/}$$

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

$$\tilde{v}_T' S_{TT}^{(1)} \tilde{v}_T = \tilde{\#}^2 \; {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \; {}_T^{(1)} + I_1 + I_2 + I_3; \tag{44}$$

where $I_1 = \hat{\#}^2 \hat{\;}_T^{(1)'} S_{TT}^{(1)-1} \hat{\;}_T^{(1)} \quad \tilde{\#}^2 \; {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \; {}_T^{(1)}$, $I_2 = \; {}_n^2 \mathrm{sgn}(\; {}_T^{(1)})' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1}$ $\hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\; {}_T^{(1)})$, $I_3 = \quad 2\hat{\#} \; {}_n \hat{\;}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\; {}_T^{(1)})$. For the term I_1, since jI_1j $\hat{\#}^2 j \hat{\;}_T^{(1)'} S_{TT}^{(1)-1} \hat{\;}_T^{(1)} \quad {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \; {}_T^{(1)} j + j\hat{\#} \quad \tilde{\#} j \; (2j\hat{\#}j + j\hat{\#} \quad \tilde{\#}j) \; {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \; {}_T^{(1)}$, 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 $\quad c_{27}[(q_n s_n)^{-1} + f\log(n)g^{-1} + \exp(\quad n \; {}_1 = 12) + \exp(\quad n \; {}_2 = 12)]$,

$$jI_1 j \quad c_{28} f \; {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \; {}_T^{(1)} g^{-1} [q_n s_n = n + f \log\log(n) = ng^{1/2}] + c_{28} \; {}_n f \; {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \; {}_T^{(1)} g^{-1/2}$$
$$[(q_n s_n)^{3/2} = n + f q_n s_n \log(q_n s_n) = ng^{1/2} + f q_n s_n \log\log(n) = ng^{1/2}]:$$

To bound the term I_2, since $jI_2 j \lesssim \; {}_n^2 q_n s_n \quad 1 + \quad f\mathrm{sgn}(\; {}_T^{(1)})' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2}$ $\mathrm{sgn}(\; {}_T^{(1)}) g \quad f\mathrm{sgn}(\; {}_T^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; {}_T^{(1)}) g^{-1} \quad 1$, it follows from Lemma 9 that there exist universal constants $c_{29}; c_{30} > 0$ such that with probability at least 1 $\; c_{29}[(q_n s_n)^{-1} +$ $f\log(n)g^{-1} + \exp(\quad n \; {}_1 = 12) + \exp(\quad n \; {}_2 = 12)]$,

$$jI_2 j \quad c_{30} \; {}_n^2 q_n s_n :$$

For the term I_3, since $jI_3 j \quad 2 \; {}_n j\#j \; j \hat{\;}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\; {}_T^{(1)}) \quad {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2}$ $\mathrm{sgn}(\; {}_T^{(1)}) j + 2j\hat{\#}j \quad f \; {}_n \; {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; {}_T^{(1)}) g$, it follows from Lemma 10, (93) and (31) that there exist constants $c_{31}; c_{32} > 0$ such that with probability at least 1 $\quad c_{31}[(q_n s_n)^{-1} +$ $f\log(n)g^{-1} + \exp(\quad n \; {}_1 = 12) + \exp(\quad n \; {}_2 = 12)]$,

$$jI_3 j \quad c_{32}(\; {}_n^2 q_n s_n)^{1/2} f \; {}_T^{(1)'} \Sigma_{TT}^{(1)-1} \; {}_T^{(1)} g^{-1/2} :$$

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 $\quad c_{33}[(q_n s_n)^{-1} + f\log(n)g^{-1} +$

$\exp(-n\kappa_1/12)+\exp(-n\kappa_2/12)]$,

$$|\tilde{v}_T' S_{TT}^{(1)}\tilde{v}_T - \tilde{\omega}^2 \beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}| \le c_{34}(\lambda_n^2 q_n s_n)^{1/2}\{\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\}^{-1/2}. \tag{45}$$

Since $|\tilde{v}_T' S_{TT}^{(1)}\tilde{v}_T - \tilde{v}_T'\Sigma_{TT}^{(1)}\tilde{v}_T| \le \lambda_{\max}(\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\|\Lambda_T^{(1)-1/2}(\Sigma_{TT}^{(1)}-S_{TT}^{(1)})\Lambda_T^{(1)-1/2}\|_2\tilde{v}_T' S_{TT}^{(1)}\tilde{v}_T$, it follows from Lemma 7 and Lemma 8 that there exist constants $c_{35}, c_{36} > 0$ such that with probability at least $1 - c_{35}\{(q_n s_n)^{-1}+\exp(-n\kappa_1/12)+\exp(-n\kappa_2/12)\}$,

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

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

$$\tilde{v}_T'\hat{\beta}_T^{(1)} = \tilde{\omega}\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)} + V_1 + V_2, \tag{47}$$

where $V_1 = \hat{\omega}\hat{\beta}_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\beta}_T^{(1)} - \tilde{\omega}\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}$ and $V_2 = -\lambda_n\hat{\beta}_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})$. Since $|V_1| \le |\hat{\omega}||\hat{\beta}_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\beta}_T^{(1)} - \beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}| + |\hat{\omega}-\tilde{\omega}|\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}$, 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 $1 - c_{37}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\kappa_1/12)+\exp(-n\kappa_2/12)]$,

$$|V_1| \le c_{38}[q_n s_n/n + \{\log\log(n)/n\}^{1/2}] + c_{38}\lambda_n\{\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\}^{1/2}$$
$$[(q_n s_n)^{3/2}/n + \{q_n s_n\log(q_n s_n)/n\}^{1/2} + \{q_n s_n\log\log(n)/n\}^{1/2}].$$

Since $|V_2| \le \lambda_n|\hat{\beta}_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) - \beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})| + \lambda_n\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})$, it holds from Lemma 10, (93), and (41) that there exist constants $c_{39}, c_{40} > 0$ such that with probability at least $1 - c_{39}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\kappa_1/12)+\exp(-n\kappa_2/12)]$,

$$|V_2| \le c_{40}(\lambda_n^2 q_n s_n)^{1/2}\{\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\}^{1/2}.$$

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 $1 - c_{41}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\kappa_1/12)+\exp(-n\kappa_2/12)]$,

$$|\tilde{v}_T'\hat{\beta}_T^{(1)} - \tilde{\omega}\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}| \le c_{42}(\lambda_n^2 q_n s_n)^{1/2}\{\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\}^{1/2}. \tag{48}$$

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

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

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

$$|\tilde{v}_T'(\hat{\ }_{1,T}^{(1)} - \,_{1,T}^{(1)})| \le \|\Lambda_T^{(1)-1/2}(\hat{\ }_{1,T}^{(1)} - \,_{1,T}^{(1)})\|_\infty \, q_n s_n \,_{\max}(\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2})^{1/2}$$
$$|\hat{\omega}| \, \{\hat{\ }_T^{(1)\prime} S_{TT}^{(1)-1} \hat{\ }_T^{(1)}\}^{1/2} + \,_n\{\mathrm{sgn}(\,_T^{(1)})' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\,_T^{(1)})\}^{1/2} .$$

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 $1 - c_{43}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\,_1/12) + \exp(-n\,_2/12)]$,

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

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

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

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

$$P\left\{\|\Lambda_T^{(1)-1/2}(\hat{\ }_{1,T}^{(1)} - \,_{1,T}^{(1)})\|_\infty \ge t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right\} \le 2q_n s_n \exp\{-(\,_1/4)nt^2\}.$$

Plugging $t = c_{45}\{\log(q_n s_n)/n\}^{1/2}$ with $c_{45} = (8/\,_1)^{1/2}$ into the above yields $P\{\|\Lambda_T^{(1)-1/2}(\hat{\ }_{1,T}^{(1)} - \,_{1,T}^{(1)})\|_\infty \le c_{45}\{\log(q_n s_n)/n\}^{1/2} \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\} \ge 1 - 2(q_n s_n)^{-1}$. Together with Lemma 3, it can be deduced that $P\{\|\Lambda_T^{(1)-1/2}(\hat{\ }_{1,T}^{(1)} - \,_{1,T}^{(1)})\|_\infty \le c_{45}\{\log(q_n s_n)/n\}^{1/2}\} \ge 1 - 2\{(q_n s_n)^{-1} + \exp(-n\,_1/12) + \exp(-n\,_2/12)\}$. Together with (50), there exist universal constants $c_{46}, c_{47} > 0$ such that with probability at least $1 - c_{46}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\,_1/12) + \exp(-n\,_2/12)]$,

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

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

$$\Omega_1 = \Omega_1^* \overset{\eta}{!} 1; \qquad \Omega_1^* \; ! \; 0: \tag{51}$$

Similar argument leads to $\Omega_2 = \Omega_2^* \overset{\eta}{!} 1$, $\Omega_2^* \; ! \; 0$. Together with (38), (39), and (51), it holds that $R^\diamond(\hat{v}) = R^\circ(\;^{(1)}) \overset{\eta}{!} 1$, $R^\circ(\;^{(1)}) \; ! \; 0$, 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) $c_1 \quad {}_{\min}(\Lambda^{\dagger 1/2}\Sigma\Lambda^{\dagger 1/2}) \quad {}_{\max}(\Lambda^{\dagger 1/2}\Sigma\Lambda^{\dagger 1/2}) \quad c_2$,

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

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

(b) $\mathrm{P}_{j \in T^*} \mathrm{P}_{k=s_n+1}^{\infty} \; !_{jk} \;\; {}_{jk}^{*2} = o(\min_{j \in T^*} \mathrm{P}_{k=1}^{s_n} \; !_{jk} \;\; {}_{jk}^{*2})$.

Then we have the following properties:

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

2) $\Delta^{(1)2} = f1 + o(r_n^{-1}) + o(r_n^{-1/2} \;\; {}_n^{1/2})g\Delta^2$,

where the parameter ${}_n = (\;{}_T^{*(1)\prime}\Sigma_{TT}^{(1,2)}\Sigma_{TT}^{(2)\dagger}\Sigma_{TT}^{(2,1)} \;\; {}_T^{*(1)}) = (\;{}_T^{*(1)\prime}\Sigma_{TT}^{(1)} \;\; {}_T^{*(1)}) \quad 1$.

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

$$\Sigma^{(1)}_{TT} \quad \Sigma^{(1)}_{TN} \quad \Sigma^{(1,2)}_{TT} \quad \Sigma^{(1,2)}_{TN}$$

Together with the triangle inequality, we have

$$
\begin{aligned}
& k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \;{}_N^{*(1)}k_2 \\
& k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)} \;{}_T^{*(2)}k_2 + k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)} \;{}_T^{*(2)}k_2 + \\
& k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)} \;{}_N^{*(2)}k_2 + k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)} \;{}_N^{*(2)}k_2;
\end{aligned}
$$

which further implies that

$$
\begin{aligned}
& k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \;{}_N^{*(1)}k_2^2 \\
\lesssim & k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)} \;{}_T^{*(2)}k_2^2 + k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)} \;{}_T^{*(2)}k_2^2 + \\
& k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)} \;{}_N^{*(2)}k_2^2 + k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)} \;{}_N^{*(2)}k_2^2: \qquad (54)
\end{aligned}
$$

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

$$
\begin{aligned}
& {}_{\min} \; \Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2} \\
= & {}_{\max}^{-1} \; \Lambda_N^{(1)1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})^{-1}\Lambda_N^{(1)1/2} \qquad c_1: \qquad (55)
\end{aligned}
$$

for the universal constant $c_1 > 0$ defined in condition (a). Hence, for the term $k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \;{}_N^{*(1)}k_2^2$, we have

$$
\begin{aligned}
& k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \;{}_N^{*(1)}k_2^2 \\
& c_1 \;{}_{\max} \; \Lambda_N^{(1)1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})^{-1}\Lambda_N^{(1)1/2} \qquad f(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \;{}_N^{*(1)}g' \\
& \quad \Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2}f(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \;{}_N^{*(1)}g \\
& c_1(\Lambda_N^{(1)1/2} \;{}_N^{*(1)})' f\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}g(\Lambda_N^{(1)1/2} \;{}_N^{*(1)}) \\
& c_1^2 k\Lambda_N^{(1)1/2} \;{}_N^{*(1)}k_2^2; \qquad (56)
\end{aligned}
$$

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

$$
\begin{aligned}
&{}_{\min}\ \Lambda_N^{(1)1/2}\Sigma_{NN}^{(1)-1}\Lambda_N^{(1)1/2} \;=\; {}^{-1}_{\max}\ \Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1)}\Lambda_N^{(1)-1/2} \quad c_2^{-1}; && (57)\\
&{}_{\min}\ \Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2} \quad c_1; && (58)\\
&{}_{\max}\ \Lambda_T^{(1)-1/2}\Sigma_{TN}^{(1)}\Sigma_{NN}^{(1)-1}\Sigma_{NT}^{(1)}\Lambda_T^{(1)-1/2} \qquad {}_{\max}\ \Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2} \quad c_2; && (59)\\
&{}_{\max}\ \Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\Lambda_T^{(2)\dagger 1/2} \qquad {}_{\max}\ \Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2)}\Lambda_T^{(2)\dagger 1/2} \quad c_2; && (60)
\end{aligned}
$$

for the universal constants c_1 and c_2 defined in condition (a). Thus, for the term $k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)}k_2^2$, we have

$$
\begin{aligned}
& k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)}k_2^2\\
& c_2\ {}_{\min}\ \Lambda_N^{(1)1/2}\Sigma_{NN}^{(1)-1}\Lambda_N^{(1)1/2}\ (\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})\\
& c_2(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})'\Sigma_{NN}^{(1)-1}(\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})\\
& c_2\ {}_{\max}\ \Lambda_T^{(1)-1/2}\Sigma_{TN}^{(1)}\Sigma_{NN}^{(1)-1}\Sigma_{NT}^{(1)}\Lambda_T^{(1)-1/2}\ k\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)}k_2^2\\
& c_2^2(\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})'(\Lambda_T^{(1)1/2}\Lambda_T^{(1)1/2})(\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})\\
& c_2^2c_1^{-1}\ {}_{\min}\ \Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}\ (\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})'(\Lambda_T^{(1)1/2}\Lambda_T^{(1)1/2})(\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\ {}_T^{*(2)})\\
& c_2^2c_1^{-1}\ {}_{\max}\ \Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)}\Lambda_T^{(2)\dagger 1/2}\ k\Lambda_T^{(2)1/2}\ {}_T^{*(2)}k_2^2\\
& c_2^3c_1^{-1}k\Lambda_T^{(2)1/2}\ {}_T^{*(2)}k_2^2;
\end{aligned}
$$

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 $k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)}\ {}_T^{*(2)}k_2^2$, we have

$$
k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)}\ {}_T^{*(2)}k_2^2 \qquad c_2^2k\Lambda_T^{(2)1/2}\ {}_T^{*(2)}k_2^2.
$$

In a similar fashion, for the term $k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)}\ {}_N^{*(2)}k_2^2$, we have

$$
k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)}\ {}_N^{*(2)}k_2^2 \qquad c_2^3c_1^{-1}k\Lambda_N^{(2)1/2}\ {}_N^{*(2)}k_2^2.
$$

In addition, for the term $k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}\ {}_N^{*(2)}k_2^2$, one has

$$k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}\ {}_N^{*(2)}k_2^2 = (\Sigma_{NN}^{(1,2)}\ {}_N^{*(2)})'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NN}^{(1,2)}\ {}_N^{*(2)})$$

$$c_2\ {}_{\min}\ \Lambda_N^{(1)1/2}\Sigma_{NN}^{(1)-1}\Lambda_N^{(1)1/2}\ \ (\Sigma_{NN}^{(1,2)}\ {}_N^{*(2)})'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NN}^{(1,2)}\ {}_N^{*(2)})$$

$$c_2(\Lambda_N^{(2)1/2}\ {}_N^{*(2)})'(\Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2,1)}\Sigma_{NN}^{(1)-1}\Sigma_{NN}^{(1,2)}\Lambda_N^{(2)\dagger 1/2})(\Lambda_N^{(2)1/2}\ {}_N^{*(2)})$$

$$c_2\ {}_{\max}\ \Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2,1)}\Sigma_{NN}^{(1)-1}\Sigma_{NN}^{(1,2)}\Lambda_N^{(2)\dagger 1/2}\ \ k\Lambda_N^{(2)1/2}\ {}_N^{*(2)}k_2^2$$

$$c_2\ {}_{\max}\ \Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2)}\Lambda_N^{(2)\dagger 1/2}\ \ k\Lambda_N^{(2)1/2}\ {}_N^{*(2)}k_2^2$$

$$c\ ^{(1)}\qquad ^{247\ 9411}$$

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

$$
\begin{aligned}
& \ _T^{*(2)'}\Sigma_{TT}^{(2)} \ _T^{*(2)} \quad \max \ \Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2)}\Lambda_T^{(2)\dagger 1/2} \ k\Lambda_T^{(2)1/2} \ _T^{*(2)}k_2^2 \\
& c_2 \sum_{j\in T^*}\sum_{k=s_n+1}^{\infty} !_{jk} \ _{jk}^{*2} \quad c_2 o(\min_{j\in T^*}\sum_{k=1}^{\infty} !_{jk} \ _{jk}^{*2}) \quad c_2 r_n^{-1} o(\sum_{j\in T^*}\sum_{k=1}^{\infty} !_{jk} \ _{jk}^{*2}) \\
& c_2 r_n^{-1} o(\sum_{j\in T^*}\sum_{k=1}^{\infty} !_{jk} \ _{jk}^{*2}) \quad \min \ \Lambda_{T^*}^{\dagger 1/2}\Sigma_{T^*T^*}\Lambda_{T^*}^{\dagger 1/2} \ (\ _{T^*}^{*'}\Lambda_{T^*}^{1/2}\Lambda_{T^*}^{1/2} \ _{T^*}^{*})o(r_n^{-1}) \\
& (\ _{T^*}^{*'}\Sigma_{T^*T^*} \ _{T^*}^{*})o(r_n^{-1}) \quad \Delta^2 o(r_n^{-1}); \qquad (63)
\end{aligned}
$$

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

$$
j \ _T^{*(1)'}\Sigma_{TT}^{(1,2)} \ _T^{*(2)}j \quad k\Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)} \ _T^{*(1)}k_2 k\Lambda_T^{(2)1/2} \ _T^{*(2)}k_2:
$$

For the term $k\Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)} \ _T^{*(1)}k_2$, we have

$$
\begin{aligned}
& k\Lambda_T^{(2)\dagger 1/2}\Sigma_{TT}^{(2,1)} \ _T^{*(1)}k_2^2 \lesssim \ _T^{*(1)'}(\Sigma_{TT}^{(1,2)}\Sigma_{TT}^{(2)\dagger}\Sigma_{TT}^{(2,1)}) \ _T^{*(1)} \lesssim \ _n \ _T^{*(1)'}\Sigma_{TT}^{(1)} \ _T^{*(1)} \\
& \lesssim \ _n k\Lambda_{T^*}^{1/2} \ _{T^*}^{*}k_2^2 \lesssim \ _n \ _{T^*}^{*'}\Sigma_{T^*T^*} \ _{T^*}^{*} \lesssim \ _n\Delta^2;
\end{aligned}
$$

where the last inequality is by (62). For the term $k\Lambda_T^{(2)1/2} \ _T^{*(2)}k_2$, one has

$$
k\Lambda_T^{(2)1/2} \ _T^{*(2)}k_2^2 \lesssim k\Lambda_{T^*}^{(2)1/2} \ _{T^*}^{*(2)}k_2^2 \lesssim o(\min_{j\in T^*}\sum_{k=1}^{\infty} !_{jk} \ _{jk}^{*2}) \lesssim \sum_{j\in T^*}\sum_{k=1}^{\infty} !_{jk} \ _{jk}^{*2} o(r_n^{-1}) \lesssim \Delta^2 o(r_n^{-1}):
$$

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

$$
j \ _T^{*(1)'}\Sigma_{TT}^{(1,2)} \ _T^{*(2)}j \lesssim \Delta^2 o(r_n^{-1/2} \ _n^{1/2}): \qquad (64)
$$

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

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

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 $\Delta^{(1)2} = f1 + o(r_n^{-1}) + o(r_n^{-1/2} \;\; {}_n^{1/2})g\Delta^2$, which completes the proof. □

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

$$\min_{v_T \in \mathbb{R}^{q_n s_n}} \text{h} \frac{1}{2} v_T' \;\; S_{TT}^{(1)} + \frac{n_1 n_2}{n(n \quad 2)} \hat{}_T^{(1)} \hat{}_T^{(1)\prime} \;\; v_T \quad \frac{n_1 n_2}{n(n \quad 2)} v_T' \hat{}_T^{(1)} + \;\; {}_n(\hat{\Lambda}_T^{(1)1/2} v_T)' sgn(\;\; {}_T^{(1)}) \text{ i};$$

where $v_T = (v_1'; : : : ; v_{q_n}')'$ with sub-vectors $v_j = (v_{j1}; : : : ; v_{js_n})'$ 2 $\mathbb{R}^{s_n}$. Let $\tilde{v}_T$ be the solution of this optimization problem where $\tilde{v}_T = (\tilde{v}_1'; : : : ; \tilde{v}_{q_n}')'$ with sub-vectors $\tilde{v}_j = (\tilde{v}_{j1}; : : : ; \tilde{v}_{js_n})'$ 2 $\mathbb{R}^{s_n}$, then we have:

$$\tilde{v}_T = f n_1 n_2 n^{-1}(n \quad 2)^{-1} g f 1 + \;\; {}_n \hat{}_T^{(1)\prime} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} sgn(\;\; {}_T^{(1)}) g$$

$$1 + f n_1 n_2 n^{-1}(n \quad 2)^{-1} g \hat{}_T^{(1)\prime} S_{TT}^{(1)-1} \hat{}_T^{(1)} \;\; {}^{-1} S_{TT}^{(1)-1} \hat{}_T^{(1)} \qquad {}_n S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} sgn(\;\; {}_T^{(1)}):$$

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

Lemma 3. *De ne the events M_n and M_n^* as*

$$M_n = f \;\; {}_1{=}2 \quad n_1{=}n \quad 3 \;\; {}_1{=}2 g \setminus f \;\; {}_2{=}2 \quad n_2{=}n \quad 3 \;\; {}_2{=}2 g;$$

$$M_n^* = f \;\; {}_1 \;\; {}_2{=}4 \quad n_1 n_2{=}n^2 \quad 9 \;\; {}_1 \;\; {}_2{=}4 g:$$

Then we have the following properties:

1) $P(M_n) \quad 1 \quad 2\exp(\;\; n \;\; {}_1{=}12) \quad 2\exp(\;\; n \;\; {}_2{=}12)$.

2) $P(M_n^*) \quad 1 \quad 2\exp(\;\; n \;\; {}_1{=}12) \quad 2\exp(\;\; n \;\; {}_2{=}12)$.

Proof of Lemma 3: First of all, note that $n_1 \quad$ Binomial$(n; \;\; {}_1)$. Invoking the chernoff tail bounds for binomial random variables, we have that for any $\;\; 2\; [0; 1]$,

$$P f n_1 \quad (1 + \;\;) n \;\; {}_1 g \quad \exp(\;\; n \;\; {}_1 \;\; {}^2{=}3);$$

$$P f n_1 \quad (1 \quad\;) n \;\; {}_1 g \quad \exp(\;\; n \;\; {}_1 \;\; {}^2{=}3):$$

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

$$P(n_1/n \quad 3 \quad _1/2) \quad \exp(\quad n \quad _1/12);$$
$$P(n_1/n \quad _1/2) \quad \exp(\quad n \quad _1/12). \tag{65}$$

Accordingly, we have

$$P(\quad _1/2 \quad n_1/n \quad 3 \quad _1/2) = 1 \quad P(n_1/n > 3 \quad _1/2) \quad P(n_1/n < \quad _1/2)$$
$$1 \quad 2\exp(\quad n \quad _1/12);$$

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

$$P(\quad _2/2 \quad n_2/n \quad 3 \quad _2/2) \quad 1 \quad 2\exp(\quad n \quad _2/12).$$

To this end, based on the above two inequalities, we can deduce that $P(M_n) \quad 1$ $2\exp(\quad n \quad _1/12) \quad 2\exp(\quad n \quad _2/12)$, which completes the proof of 1). Property 2) follows from the fact that $M_n \quad M_n^*$. □

Lemma 4. *For any % 2* $(e^{-n/100}; 1/100)$*, de ne the event* $M_{3n}(\%)$ *as*

$$M_{3n}(\%) = \Big\{ \hat{\ }^{(1)\prime}_T S^{(1)-1}_{TT} \hat{\ }^{(1)}_T \quad ^{(1)\prime}_T \Sigma^{(1)-1}_{TT} \ ^{(1)}_T \lesssim q_n s_n/n + \log(\%^{-1})/n$$
$$+ \quad q_n s_n/n + \{\log(\%^{-1})/n\}^{1/2} \{ \ ^{(1)\prime}_T \Sigma^{(1)-1}_{TT} \ ^{(1)}_T$$

Proof of Lemma 4: First of all, note that

$$
\hat{}_{T}^{(1)\prime} S_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)}
$$

$$
=(\hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)})(\hat{}_{T}^{(1)\prime} S_{TT}^{(1)-1} \hat{}_{T}^{(1)} = \hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad 1)
$$

$$
+ \ {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} (\hat{}_{T}^{(1)\prime} S_{TT}^{(1)-1} \hat{}_{T}^{(1)} = \hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad 1) + (\hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)});
$$

which implies that

$$
j \hat{}_{T}^{(1)\prime} S_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} j
$$

$$
j \hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} j \ \ j \hat{}_{T}^{(1)\prime} S_{TT}^{(1)-1} \hat{}_{T}^{(1)} = \hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad 1 j
$$

$$
+ \ {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} j \hat{}_{T}^{(1)\prime} S_{TT}^{(1)-1} \hat{}_{T}^{(1)} = \hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad 1 j + j \hat{}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} j:
$$

Together with Lemma 18 and Lemma 19, we conclude that with probability at least 1 4% 4 exp(n $_1$=12) 4 exp(n $_2$=12),

$$
j \hat{}_{T}^{(1)\prime} S_{TT}^{(1)-1} \hat{}_{T}^{(1)} \quad {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} j \lesssim q_n s_n = n + \log(\%^{-1}) = n + \ q_n s_n = n + f \log(\%^{-1}) = n g^{1/2}
$$

$$
f \ {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} g + f \log(\%^{-1}) = n g^{1/2} f \ {}_{T}^{(1)\prime} \Sigma_{TT}^{(1)-1} {}_{T}^{(1)} g^{1/2};
$$

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) $P \ k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \quad I_{q_n s_n} k_{\max} \quad c_1 f \log(q_n s_n) = n g^{1/2} \quad 1 \quad c_2 f (q_n s_n)^{-1}$

$+ \exp(\ n \ _1 = 12) + \exp(\ n \ _2 = 12) g.$

2) $P \ k \Lambda_T^{(1)} \hat{\Lambda}_T^{(1)-1} \quad I_{q_n s_n} k_{\max} \quad c_1 f \log(q_n s_n) = n g^{1/2} \quad 1 \quad c_2 f (q_n s_n)^{-1}$

$+ \exp(\ n \ _1 = 12) + \exp(\ n \ _2 = 12) g.$

3) $P\left\{\|\hat\Lambda_T^{(1)1/2}\Lambda_T^{(1)-1/2} - I_{q_ns_n}\|_{\max} \le c_1\{\log(q_ns_n)/n\}^{1/2}\right\} \ge 1 - c_2\{(q_ns_n)^{-1}$

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

4) $P\left\{\|\Lambda_T^{(1)1/2}\hat\Lambda_T^{(1)-1/2} - I_{q_ns_n}\|_{\max} \le c_1\{\log(q_ns_n)/n\}^{1/2}\right\} \ge 1 - c_2\{(q_ns_n)^{-1}$

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

Note that $I_{q_ns_n}$ denotes the $q_ns_n \times q_ns_n$ 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 $\|X\|_{\psi_1} = \sup_{q\ge 1} q^{-1}\{E(|X|^q)\}^{1/q}$. Now, we are in a position to start the proof. First of all, notice that

$$\|\hat\Lambda_T^{(1)}\Lambda_T^{(1)-1} - I_{q_ns_n}\|_{\max} = \max_{j\in T}\max_{k\le s_n}|\hat\omega_{jk}\omega_{jk}^{-1} - 1|. \tag{66}$$

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

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

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

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

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

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

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

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

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

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 $\{\omega_{jk}^{-1/2}(\epsilon_{ijk} - \mu_{1jk})\}^2$ satisfy

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

For the term Υ_1, conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, one can show that for any $t \ge 0$,

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

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 $t \ge 0$,

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

for some universal constant $c_2 > 0$. Regarding the term Υ_3, it is clear that for any $t \geq 0$,

$$
\begin{aligned}
& P\left(\Upsilon_3 \geq t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) \\
\leq & \sum_{j\in T}\sum_{k\leq s_n} P\left[\left| n_1^{-1}\sum_{i_1\in H_1} \omega_{jk}^{-1/2}(\epsilon_{i_1jk} - \bar{\epsilon}_{1jk})\right| \geq t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right] \\
\leq & \, 2q_n s_n \exp(-c_3 n t^2),
\end{aligned} \tag{71}
$$

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 $t \geq 0$,

$$
P\left(\Upsilon_4 \geq t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) \leq 2q_n s_n \exp(-c_4 n t^2), \tag{72}
$$

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 $t \geq 0$,

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

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

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

It then follows that

$$\begin{aligned}
&P\left(\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le c_7\{\log(q_n s_n)/n\}^{1/2}\right) \\
&\ge \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\left(\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le c_7\{\log(q_n s_n)/n\}^{1/2} \mid \{Y_i = y_i\}_{i=1}^n\right) P\left(\{Y_i = y_i\}_{i=1}^n\right) \\
&\ge \{1 - 8(q_n s_n)^{-1}\} \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\left(\{Y_i = y_i\}_{i=1}^n\right) = \{1 - 8(q_n s_n)^{-1}\} P(\mathcal{M}_n) \\
&\ge 1 - 8\{(q_n s_n)^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)\},
\end{aligned}$$

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 $\left(\|\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le c_7\{\log(q_n s_n)/n\}^{1/2}\right)$,

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

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

$$\begin{aligned}
&P\left(\|\Lambda_T^{(1)}\hat{\Lambda}_T^{(1)-1} - I_{q_n s_n}\|_{\max} \le 2c_7\{\log(q_n s_n)/n\}^{1/2}\right) \\
&\ge 1 - 8\{(q_n s_n)^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)\},
\end{aligned} \tag{74}$$

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) $P\left(\|\hat{\Lambda}_N^{(1)}\Lambda_N^{(1)-1} - I_{(p_n-q_n)s_n}\|_{\max} \le c_1[\log\{(p_n - q_n)s_n\}/n]^{1/2}\right) \ge 1 - c_2[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\varepsilon_1/12) + \exp(-n\varepsilon_2/12)]$.

2) $P\left(\|\Lambda_N^{(1)}\hat{\Lambda}_N^{(1)-1} - I_{(p_n-q_n)s_n}\|_{\max} \le c_1[\log\{(p_n - q_n)s_n\}/n]^{1/2}\right) \ge 1 - c_2[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\varepsilon_1/12) + \exp(-n\varepsilon_2/12)]$.

3) $P\left(\|\hat{\Lambda}_N^{(1)1/2}\Lambda_N^{(1)-1/2} - I_{(p_n-q_n)s_n}\|_{\max} \le c_1[\log\{(p_n - q_n)s_n\}/n]^{1/2}\right) \ge 1 - c_2[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\varepsilon_1/12) + \exp(-n\varepsilon_2/12)]$.

4) $P\left(\|\Lambda_N^{(1)1/2}\hat{\Lambda}_N^{(1)-1/2} - I_{(p_n-q_n)s_n}\|_{\max} \le c_1[\log\{(p_n - q_n)s_n\}/n]^{1/2}\right) \ge 1 - c_2[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\varepsilon_1/12) + \exp(-n\varepsilon_2/12)]$.

5) $P\{\det(\hat{\Lambda}_N^{(1)}) \neq 0\} \ge 1 - c_2[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\varepsilon_1/12) + \exp(-n\varepsilon_2/12)]$.

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 $1 - c_3[\{(p_n - q_n)s_n\}^{-1} + \exp(-n\varepsilon_1/12) + \exp(-n\varepsilon_2/12)]$,

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

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:

$$P\left\{\|\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} - \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}\|_2 \le c_1 \{q_n^2 s_n^2 \log(q_n s_n)/n\}^{1/2}\right\}$$
$$\ge 1 - c_2\{(q_n s_n)^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)\}.$$

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

$$\begin{aligned}
\Omega_1 =& 2n^{-1} n_1 q_n s_n \max_{j_2 \in T}\max_{k_2 \le s_n}\max_{j_1 \in T}\max_{k_1 \le s_n} \Big| n_1^{-1} \sum_{i \in H_1} \Big[\omega_{j_1k_1}^{-1/2}(\xi_{ij_1k_1} - \mu_{1j_1k_1}) \\
& \omega_{j_2k_2}^{-1/2}(\xi_{ij_2k_2} - \mu_{1j_2k_2}) - \mathrm{corr}(\xi_{j_1k_1}, \xi_{j_2k_2})\Big]\Big|, \\
\Omega_2 =& 2n^{-1} n_2 q_n s_n \max_{j_2 \in T}\max_{k_2 \le s_n}\max_{j_1 \in T}\max_{k_1 \le s_n} \Big| n_2^{-1} \sum_{i \in H_2} \Big[\omega_{j_1k_1}^{-1/2}(\xi_{ij_1k_1} - \mu_{2j_1k_1}) \\
& \omega_{j_2k_2}^{-1/2}(\xi_{ij_2k_2} - \mu_{2j_2k_2}) - \mathrm{corr}(\xi_{j_1k_1}, \xi_{j_2k_2})\Big]\Big|, \\
\Omega_3 =& 2n^{-1} n_1 q_n s_n \max_{j_2 \in T}\max_{k_2 \le s_n}\max_{j_1 \in T}\max_{k_1 \le s_n} \Big| n_1^{-1} \sum_{i_1 \in H_1} \omega_{j_1k_1}^{-1/2}(\xi_{i_1j_1k_1} - \mu_{1j_1k_1}) \\
& n_1^{-1} \sum_{i_1 \in H_1} \omega_{j_2k_2}^{-1/2}(\xi_{i_1j_2k_2} - \mu_{1j_2k_2}) \Big|, \\
\Omega_4 =& 2n^{-1} n_2 q_n s_n \max_{j_2 \in T}\max_{k_2 \le s_n}\max_{j_1 \in T}\max_{k_1 \le s_n} \Big| n_2^{-1} \sum_{i_2 \in H_2} \omega_{j_1k_1}^{-1/2}(\xi_{i_2j_1k_1} - \mu_{2j_1k_1}) \\
& n_2^{-1} \sum_{i_2 \in H_2} \omega_{j_2k_2}^{-1/2}(\xi_{i_2j_2k_2} - \mu_{2j_2k_2}) \Big|, \\
\Omega_5 =& 4n^{-1} q_n s_n.
\end{aligned}$$

For the term Ω_1, conditional on any nonempty $\{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$, it can be shown that for any $t \geq 0$,

$$
\begin{aligned}
& P\left(\Omega_1 \geq t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) \\
\leq\ & P\Bigg(\max_{j_2\in T}\max_{k_2\leq s_n}\max_{j_1\in T}\max_{k_1\leq s_n} \Big| n_1^{-1} \sum_{i\in H_1} \Big[\omega_{j_1k_1}^{-1/2}(\epsilon_{ij_1k_1} - \mu_{1j_1k_1}) \omega_{j_2k_2}^{-1/2}(\epsilon_{ij_2k_2} - \mu_{1j_2k_2}) \\
& \quad - \mathrm{corr}(\epsilon_{j_1k_1}, \epsilon_{j_2k_2}) \Big] \Big| \geq (3\lambda_1 q_n s_n)^{-1} t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \Bigg) \\
\leq\ & \sum_{j_2\in T}\sum_{k_2=1}^{s_n}\sum_{j_1\in T}\sum_{k_1=1}^{s_n} P\Bigg(\Big| n_1^{-1} \sum_{i\in H_1} \Big[\omega_{j_1k_1}^{-1/2}(\epsilon_{ij_1k_1} - \mu_{1j_1k_1}) \omega_{j_2k_2}^{-1/2}(\epsilon_{ij_2k_2} - \mu_{1j_2k_2}) \\
& \quad - \mathrm{corr}(\epsilon_{j_1k_1}, \epsilon_{j_2k_2}) \Big] \Big| \geq (3\lambda_1 q_n s_n)^{-1} t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \Bigg) \\
\leq\ & \sum_{j_2\in T}\sum_{k_2=1}^{s_n}\sum_{j_1\in T}\sum_{k_1=1}^{s_n} 2\exp\left(-c_1 n \min\{(q_n s_n)^{-2} t^2, (q_n s_n)^{-1} t\} \right) \\
=\ & 2(q_n s_n)^2 \exp\left(-c_1 n \min\{(q_n s_n)^{-2} t^2, (q_n s_n)^{-1} t\} \right),
\end{aligned}
$$

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

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

Similar reasoning yields that

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

for some universal constant $c_3 > 0$. For the term Ω_3, it is apparent to see that for any

$t \geq 0,$

$$
\begin{aligned}
&P\Big(\Omega_3 \geq t \,\Big|\, \{Y_i = y_i\}_{i=1}^n \cap M_n\Big) \\
&\leq \sum_{j_2\in T}\sum_{k_2=1}\sum_{j_1\in T}\sum_{k_1=1} P\Big(\Big| n_1^{-1} \sum_{i_1\in H_1} \omega_{j_1k_1}^{-1/2}(\xi_{i_1j_1k_1} - \mu_{1j_1k_1})\Big| \geq (3\kappa_1 q_n s_n)^{-1/2} t^{1/2} \\
&\quad \Big| \{Y_i = y_i\}_{i=1}^n \cap M_n\Big) + \sum_{j_2\in T}\sum_{k_2=1}\sum_{j_1\in T}\sum_{k_1=1} P\Big(\Big| n_1^{-1} \sum_{i_1\in H_1} \omega_{j_2k_2}^{-1/2}(\xi_{i_1j_2k_2} - \mu_{1j_2k_2})\Big| \\
&\quad \geq (3\kappa_1 q_n s_n)^{-1/2} t^{1/2} \,\Big|\, \{Y_i = y_i\}_{i=1}^n \cap M_n\Big) \\
&\leq 4(q_n s_n)^2 \exp(-c_4 n q_n^{-1} s_n^{-1} t),
\end{aligned}
$$

for some universal constant $c_4 > 0$, where the last inequality follows from Hoeffding inequality and the definition of 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

$$
P\Big(\Omega_3 \geq c_5 q_n s_n \log(q_n s_n)/n \,\Big|\, \{Y_i = y_i\}_{i=1}^n \cap M_n\Big) \leq 4(q_n s_n)^{-1}.
$$

Similar reasoning leads to

$$
P\Big(\Omega_4 \geq c_6 q_n s_n \log(q_n s_n)/n \,\Big|\, \{Y_i = y_i\}_{i=1}^n \cap M_n\Big) \leq 4(q_n s_n)^{-1},
$$

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

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

Finally, we have

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

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 $1 \quad c_3 f(q_n s_n)^{-1} + \exp(\quad n \quad {}_1 = 12) + \exp(\quad n \quad {}_2 = 12) g$,

$$k\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \quad c_4 f q_n^2 s_n^2 \log(q_n s_n) = n g^{1/2} + c_2 \quad 2c_2;$$

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

$$k\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} k_2 = \quad {}_{\min}^{-1}(\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}): \tag{79}$$

Moreover, it is apparent to deduce that

$${}_{\min}(\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \quad c_1 \quad k\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \quad \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}$$

2) $P \quad fsgn({}^{(1)}_{T})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn({}^{(1)}_{T})g{=}fsgn({}^{(1)}_{T})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}sgn({}^{(1)}_{T})g$

$$1 \quad c_3 \; f\log(q_n s_n){=}ng^{1/2} + f\log\log(n){=}ng^{1/2}$$

$$1 \quad c_4 \; (q_n s_n)^{-1} + f\log(n)g^{-1} + \exp(\;\; n \;\; _1{=}12) + \exp(\;\; n \;\; _2{=}12) \; .$$

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

$$\text{sgn}({}^{(1)}_{T})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\text{sgn}({}^{(1)}_{T}) = \text{sgn}({}^{(1)}_{T})'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\text{sgn}({}^{(1)}_{T}) + \Omega_1 + 2\Omega_2; \tag{80}$$

where

$$\Omega_1 = \text{sgn}({}^{(1)}_{T})'(\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} \quad I_{q_n s_n})(\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2})(\Lambda_T^{(1)-1/2}\hat{\Lambda}_T^{(1)1/2} \quad I_{q_n s_n})\text{sgn}({}^{(1)}_{T});$$

$$\Omega_2 = \text{sgn}({}^{(1)}_{T})'(\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2})(\Lambda_T^{(1)-1/2}\hat{\Lambda}_T^{(1)1/2} \quad I_{q_n s_n})\text{sgn}({}^{(1)}_{T}):$$

For the term Ω_1, it can be deduced that

$$\Omega_1 \quad q_n s_n k\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}k_2 \;\; k\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} \quad I_{q_n s_n}k^2_{\max}:$$

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

$$P\, f\Omega_1 \quad c_3 q_n s_n \log(q_n s_n){=}ng \quad 1 \quad c_4 f(q_n s_n)^{-1} + \exp(\;\; n \;\; _1{=}12) + \exp(\;\; n \;\; _2{=}12)g: \tag{81}$$

For the term Ω_2, one has

$$j\Omega_2 j \quad k(\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\text{sgn}({}^{(1)}_{T})k_1 \;\; k(\Lambda_T^{(1)-1/2}\hat{\Lambda}_T^{(1)1/2} \quad I_{q_n s_n})\text{sgn}({}^{(1)}_{T})k_\infty$$
$$q_n s_n k\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}k_2 \;\; k\hat{\Lambda}_T^{(1)1/2}\Lambda_T^{(1)-1/2} \quad I_{q_n s_n}k_{\max}:$$

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

$$P\,[j\Omega_2 j \quad c_5 f q_n^2 s_n^2 \log(q_n s_n){=}ng^{1/2}] \quad 1 \quad c_6 f(q_n s_n)^{-1} + \exp(\;\; n \;\; _1{=}12) + \exp(\;\; n \;\; _2{=}12)g:$$

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 $1 - c_7\{(q_n s_n)^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)\}$,

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

Moreover, we note that

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

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 $1 - c_9[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)]$,

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

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

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

$$= \{\mathrm{sgn}(\ _T^{(1)})'\hat{\Lambda}_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})\}/\{\mathrm{sgn}(\ _T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})\} \quad 1$$

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

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) $c_1 \quad {}_{\min}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \quad {}_{\max}(\Lambda_T^{(1)-1/2}\Sigma_{TT}^{(1)}\Lambda_T^{(1)-1/2}) \quad c_2$, *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* $1 \quad c_3[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(\ \ n\ \ _1/12) + \exp(\ \ n\ \ _2/12)]$, *we have:*

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

$$c_4[q_n s_n/n + \{\log(q_n s_n)/n\}^{1/2} + \{\log\log(n)/n\}^{1/2}] \ \ |\ {}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})|$$

$$+ c_4(q_n s_n)^{1/2}\{\log\log(n)/n\}^{1/2}[1 + \ {}_T^{(1)'}\Sigma_{TT}^{(1)-1}\ {}_T^{(1)} + \{\ {}_T^{(1)'}\Sigma_{TT}^{(1)-1}\ {}_T^{(1)}\log\log(n)/n\}^{1/2}]^{1/2}$$

$$+ c_4(q_n s_n)^{1/2}\{\log(q_n s_n)/n\}^{1/2}\{q_n s_n\log(q_n s_n\log n)/n\}^{1/2}$$

$$[1 + \ {}_T^{(1)'}\Sigma_{TT}^{(1)-1}\ {}_T^{(1)} + \{\ {}_T^{(1)'}\Sigma_{TT}^{(1)-1}\ {}_T^{(1)}\log\log(n)/n\}^{1/2}]^{1/2}.$$

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

$$|\hat{\ }_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)}) \quad {}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})| \quad \Omega_1 + \Omega_2; \qquad (82)$$

where

$$\Omega_1 = |\hat{\ }_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)}) \quad {}_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})|;$$

$$\Omega_2 = |\hat{\ }_T^{(1)'}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)}) \quad \hat{\ }_T^{(1)'}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})|.$$

For the term Ω_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 $1 - c_3[\{\log(n)\}^{-1} + \exp(-n\pi_1/12) + \exp(-n\pi_2/12)]$,

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

For the term Ω_2, it is clear that

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

where

$$\Pi_1 = |\hat{\gamma}_T^{(1)'} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} (\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} - I_q$$

$1 \quad c_7[(q_n s_n)^{-1} + f\log(n)g^{-1} + \exp(\quad n \quad {}_1=12) + \exp(\quad n \quad {}_2=12)]$,

$$\Pi_1 \quad c_8 f\log(q_n s_n)=ng^{1/2} \quad j \; {}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})j+$$
$$c_8 f q_n s_n \log(q_n s_n)=ng^{1/2} f q_n s_n \log(q_n s_n \log n)=ng^{1/2}$$
$$[1+ \; {}_T^{(1)\prime}\Sigma_{TT}^{(1)-1} \; {}_T^{(1)} + f \; {}_T^{(1)\prime}\Sigma_{TT}^{(1)-1} \; {}_T^{(1)} \log\log(n)=ng^{1/2}]^{1/2}: \tag{85}$$

To bound the term Π_2, we note that

$$\Pi_2 \quad c_1^{-1} q_n s_n(1+\Upsilon_1) \quad j\Upsilon_2 j + fj \; {}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})j + \Upsilon_3 g \quad \Upsilon_1; \tag{86}$$

where c_1 is defined in condition (b) and

$$\Upsilon_1 = jf\mathrm{sgn}(\; {}_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})gf\mathrm{sgn}(\; {}_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})g^{-1} \quad 1j;$$
$$\Upsilon_2 = f\,\hat{}_T^{(1)\prime} S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})gf\mathrm{sgn}(\; {}_T^{(1)})'\Lambda_T^{(1)1/2} S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})g^{-1}$$
$$f\,\hat{}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})gf\mathrm{sgn}(\; {}_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})g^{-1};$$
$$\Upsilon_3 = j\,\hat{}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)}) \qquad {}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})j:$$

For the term Υ_1, Lemma 22 entails that there exist universal constants $c_9 > 0$ and $c_{10} > 0$ such that with probability at least $1 \quad c_9[f\log(n)g^{-1} + \exp(\quad n \quad {}_1=12) + \exp(\quad n \quad {}_2=12)]$,

$$\Upsilon_1 \quad c_{10}[q_n s_n = n + f\log\log(n)=ng^{1/2}]: \tag{87}$$

For the term Υ_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 $fY_i = y_i g_{i=1}^n \setminus M_n \setminus f\,\hat{}_T g$, and for any $t \quad 0$,

$$P \quad j\Upsilon_2 j \quad t \; fY_i = y_i g_{i=1}^n \setminus M_n \setminus f\,\hat{}_T g$$
$$c_{11}\exp \quad c_{12} n f\,\hat{}_T'\Sigma_{TT}^{(1)-1}\,\hat{}_T g^{-1} f\mathrm{sgn}(\; {}_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\; {}_T^{(1)})gt^2 \; :$$

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

$$\begin{aligned}
&P\Big(|\Upsilon_2| \geq c_{13}\{\hat{\beta}_T'\Sigma_{TT}^{(1)-1}\hat{\beta}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\qquad \{\log\log(n)/n\}^{1/2} \,\Big|\, \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\beta}_T\}\Big) \\
&\quad \geq 1 - c_{11}\{\log(n)\}^{-1}. \qquad (88)
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&P\Big(|\Upsilon_2| \geq c_{13}\{\hat{\beta}_T'\Sigma_{TT}^{(1)-1}\hat{\beta}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\quad \{\log\log(n)/n\}^{1/2}\Big) \\
&\geq \sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n} P\Big(|\Upsilon_2| \geq c_{13}\{\hat{\beta}_T'\Sigma_{TT}^{(1)-1}\hat{\beta}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\quad \{\log\log(n)/n\}^{1/2} \,\Big|\, \{Y_i = y_i\}_{i=1}^n\Big) P\Big(\{Y_i = y_i\}_{i=1}^n\Big) \\
&= \sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n} \Big\{\int_{\hat{\nu}_T} P\Big(|\Upsilon_2| \geq c_{13}\{\hat{\beta}_T'\Sigma_{TT}^{(1)-1}\hat{\beta}_T\}^{1/2}\{\mathrm{sgn}(\beta_T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}^{-1/2} \\
&\quad \{\log\log(n)/n\}^{1/2} \,\Big|\, \{Y_i = y_i\}_{i=1}^n \cap \{\hat{\beta}_T\}\Big) f(\hat{\beta}_T|\{Y_i = y_i\}_{i=1}^n)d\hat{\beta}_T\Big\} P\Big(\{Y_i = y_i\}_{i=1}^n\Big) \\
&\geq [1 - c_{11}\{\log(n)\}^{-1}] \sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n} P\Big(\{Y_i = y_i\}_{i=1}^n\Big) = [1 - c_{11}\{\log(n)\}^{-1}]\, P(\mathcal{M}_n) \\
&\geq 1 - c_{14}\Big[\{\log(n)\}^{-1} + \exp(-n\lambda_1/12) + \exp(-n\lambda_2/12)\Big],
\end{aligned}$$

for some universal constant $c_{14} > 0$, where $f(\hat{\beta}_T|\{Y_i = y_i\}_{i=1}^n)$ 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 $1 - c_{15}[\{\log(n)\}^{-1} + \exp(-n\lambda_1/12) + \exp(-n\lambda_2/12)]$,

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

For the term Υ_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 $1 \quad c_{17}[f\log(n)g^{-1} + \exp(\ n \ _1{=}12) + \exp(\ n \ _2{=}12)]$,

$$\Upsilon_3 \quad c_{18} f q_n s_n \log\log(n){=}ng^{1/2}{:}$$

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 $1 \quad c_{19}[f\log(n)g^{-1} + \exp(\ n \ _1{=}12) + \exp(\ n \ _2{=}12)]$,

$$\begin{aligned}
\Pi_2 \quad & c_{20}[q_n s_n{=}n + \log\log(n){=}n + \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1} \ _T^{(1)} + f \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1} \ _T^{(1)} \log\log(n){=}ng^{1/2}]^{1/2} \\
& (q_n s_n)^{1/2} f\log\log(n){=}ng^{1/2} + c_{20}[q_n s_n{=}n + f\log\log(n){=}ng^{1/2}] \ \ j \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})j \\
& + c_{20} f q_n s_n \log\log(n){=}ng^{1/2} \quad [q_n s_n{=}n + f\log\log(n){=}ng^{1/2}]{:}
\end{aligned}$$

Together with (84) and (85), there exist universal constants $c_{21} > 0$ and $c_{22} > 0$ such that with probability at least $1 \quad c_{21}[(q_n s_n)^{-1} + f\log(n)g^{-1} + \exp(\ n \ _1{=}12) + \exp(\ n \ _2{=}12)]$,

$$\begin{aligned}
\Omega_2 \quad & c_{22}(q_n s_n)^{1/2} f\log\log(n){=}ng^{1/2} \\
& [q_n s_n{=}n + \log\log(n){=}n + \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1} \ _T^{(1)} + f \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1} \ _T^{(1)} \log\log(n){=}ng^{1/2}]^{1/2} \\
& + c_{22}[q_n s_n{=}n + f\log(q_n s_n){=}ng^{1/2} + f\log\log(n){=}ng^{1/2}] \ \ j \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})j \\
& + c_{22} f q_n s_n \log\log(n){=}ng^{1/2} \quad [q_n s_n{=}n + f\log\log(n){=}ng^{1/2}] \\
& + c_{22} f q_n s_n \log(q_n s_n){=}ng^{1/2} f q_n s_n \log(q_n s_n \log n){=}ng^{1/2} \\
& [1 + \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1} \ _T^{(1)} + f \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1} \ _T^{(1)} \log\log(n){=}ng^{1/2}]^{1/2}{:}
\end{aligned}$$

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 $1 \quad c_{23}[(q_n s_n)^{-1} + f\log(n)g^{-1} +$

$\exp(\; n \; _1=12) + \exp(\; n \; _2=12)],$

$$j \,\hat{}_T^{(1)\prime} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\; _T^{(1)}) \qquad _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; _T^{(1)}) j$$

$$c_{24}[q_n s_n = n + f \log(q_n s_n) = n g^{1/2} + f \log\log(n) = n g^{1/2}] \;\; j \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; _T^{(1)}) j$$

$$+ c_{24}(q_n s_n)^{1/2} f \log\log(n) = n g^{1/2} [1 + \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} + f \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} \log\log(n) = n g^{1/2}]^{1/2}$$

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

$$[1 + \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} + f \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} \log\log(n) = n g^{1/2}]^{1/2};$$

which completes the proof. □

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

(a) $\max f q_n^2 s_n^2 \log(q_n s_n); q_n s_n \log(p_n \quad q_n) g = o(n).$

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

(c) $K_1 \log f(p_n \quad q_n) s_n \log n g = (n \; _n^2) \qquad \mathsf{P}_{j \in T} \mathsf{P}_{k=1}^{s_n} \, !_{jk} \; _{jk}^2 \qquad K_2 \log f(p_n \quad q_n) s_n \log n g = (n \; _n^2) \;\; !$ *1 , for some su ciently large universal constants* $K_2 > K_1 > 0$.

(d) $\min_{j \in T} \min_{k \le s_n} !_{jk}^{1/2} j \; _{jk} j > K_3 [\log f(p_n \quad q_n) s_n \log n g = (n \; _n^2)]^{1/2} f \log(q_n s_n \log n) = n g^{1/2} +$ $K_3 [\log f(p_n \quad q_n) s_n \log n g = (n \; _n)] \; k \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} sgn(\; _T^{(1)}) k_\infty + K_3 [\log f(p_n \quad q_n) s_n \log n g = (n \; _n)]$ $[f q_n s_n \log(q_n s_n) = n g^{1/2} + f q_n s_n \log\log(n) = n g^{1/2}]$, *for some su ciently large universal constant* $K_3 > 0$.

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

$$
\begin{aligned}
&P\{sgn(\tilde{v}_T) = sgn(\beta_T^{(1)}) = sgn(\hat{\beta}_T^{(1)})\} \\
&\geq 1 - c_3[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)];
\end{aligned}
$$

where $\hat{\beta}_T^{(1)} = S_{TT}^{(1)-1}\hat{\delta}_T^{(1)}$*, 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 S_1 and S_2 as

$$
S_1 = \{k : e_k'\Sigma_{TT}^{(1)-1}\delta_T^{(1)} > 0\}; \qquad S_2 = \{k : e_k'\Sigma_{TT}^{(1)-1}\delta_T^{(1)} < 0\}.
$$

By definition, we have $S_1 \cup S_2 = \{1, \dots, q_n s_n\}$. 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

$$
\begin{aligned}
&P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{ e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\delta}_T^{(1)} \geq e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\delta_T^{(1)} \\
&\quad - c_3[q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]\, e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\delta_T^{(1)} \\
&\quad - c_3\{\log(q_n s_n \log n)/n\}^{1/2}\{\delta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\delta_T^{(1)}\}^{1/2}\Big\}\Big] \\
&\geq 1 - c_4[\{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)].
\end{aligned} \tag{90}
$$

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

$$
\delta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\delta_T^{(1)} \leq \log\{(p_n - q_n)s_n \log n\}/(n\epsilon_n^2) \to \infty; \quad \text{as} \quad n \to \infty.
$$

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

$$
\begin{aligned}
&P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{ e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\delta}_T^{(1)} \geq c_5 e_k'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\delta_T^{(1)} \\
&\quad - c_6[\log\{(p_n - q_n)s_n \log n\}/(n\epsilon_n^2)]^{1/2}\{\log(q_n s_n \log n)/n\}^{1/2}\Big\}\Big] \\
&\geq 1 - c_7[\{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)].
\end{aligned} \tag{91}
$$

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

$$P\left[\bigcap_{k\in\mathcal{S}_1}\left\{e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\ }_T^{(1)} > 0\right\}\right] \quad 1 \quad c_7[\{\log(n)\}^{-1}+\exp(\ \ n\ \ _1/12)+\exp(\ \ n\ \ _2/12)].$$

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

$$P\left[\bigcap_{k\in\mathcal{S}_2}\left\{e_k'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\ }_T^{(1)} < 0\right\}\right] \quad 1 \quad c_8[\{\log(n)\}^{-1}+\exp(\ \ n\ \ _1/12)+\exp(\ \ n\ \ _2/12)].$$

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

$$P\{\mathrm{sgn}(\ \ _T^{(1)}) = \mathrm{sgn}(\hat{\ }_T^{(1)})\} \quad 1 \quad c_9[\{\log(n)\}^{-1}+\exp(\ \ n\ \ _1/12)+\exp(\ \ n\ \ _2/12)]. \quad (92)$$

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

$$\tilde{v}_T = \hat{\#}S_{TT}^{(1)-1}\hat{\ }_T^{(1)} \quad {}_nS_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\ \ _T^{(1)});$$

where

$$\hat{\#} = \{n_1n_2n^{-1}(n \quad 2)^{-1}\}\{1+\ \ _n\hat{\ }_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\ \ _T^{(1)})\}$$
$$1+\{n_1n_2n^{-1}(n \quad 2)^{-1}\}\hat{\ }_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\ }_T^{(1)}\ \ ^{-1}.$$

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 $1 \quad c_{10}[(q_ns_n)^{-1} + \{\log(n)\}^{-1}+\exp(\ \ n\ \ _1/12)+\exp(\ \ n\ \ _2/12)]$,

$${}_n\hat{\ }_T^{(1)\prime}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}\mathrm{sgn}(\ \ _T^{(1)}) = \ \ _n\ \ _T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ \ _T^{(1)})\{1+o(1)\}+o(1).$$

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 $1 \quad c_{11}[\{\log(n)\}^{-1} + \exp(\ \ n\ \ _1/12)+\exp(\ \ n\ \ _2/12)]$,

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

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 $1 \quad c_{12}[(q_n s_n)^{-1} + f\log(n)g^{-1} + \exp(\; n \; _1{=}12) + \exp(\; n \; _2{=}12)]$,

$$\hat{\#} \quad c_{13} \; _1 \; _2 f1 + \; _n \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; _T^{(1)}) g f 1 + \; _1 \; _2 \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} g^{-1}:$$

For the term $_n \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; _T^{(1)})$, one has

$$\begin{aligned} &_n \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; _T^{(1)}) \qquad _n f \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} g^{1/2} \\ &f \mathrm{sgn}(\; _T^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\; _T^{(1)}) g^{1/2} \lesssim \; _n f q_n s_n \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} g^{1/2} \\ &\lesssim [q_n s_n \log f(p_n \quad q_n) s_n \log n g{=}n]^{1/2} \lesssim o(1); \end{aligned} \tag{93}$$

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 $1 \quad c_{14}[(q_n s_n)^{-1} + f\log(n)g^{-1} + \exp(\; n \; _1{=}12) + \exp(\; n \; _2{=}12)]$,

$$\hat{\#} \quad c_{15} f \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} g^{-1}:$$

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

$$\begin{aligned} P \overset{h}{} \bigcap_{k \in \mathcal{S}_1} \overset{n}{} \; & \hat{\#} e_k' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\;}_T^{(1)} \quad c_{17} f \; _T^{(1)\prime} \Sigma_{TT}^{(1)-1} \; _T^{(1)} g^{-1} (e_k' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \; _T^{(1)} \\ & c_{18}[\log f(p_n \quad q_n) s_n \log n g{=}(n \; _n^2)]^{1/2} f\log(q_n s_n \log n){=}n g^{1/2}) \overset{o\,i}{} \\ & 1 \quad c_{16}[(q_n s_n)^{-1} + f\log(n)g^{-1} + \exp(\; n \; _1{=}12) + \exp(\; n \; _2{=}12)]: \end{aligned}$$

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

$$
\begin{aligned}
P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{ & \lambda_n |e_k' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| \le \lambda_n |e_k' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| \\
& + c_{19}\lambda_n [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]\, |e_k' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| \\
& + c_{19}\lambda_n [\{q_n s_n \log(q_n s_n)/n\}^{1/2} + \{q_n s_n \log\log(n)/n\}^{1/2}] \Big\}\Big] \\
\ge\ & 1 - c_{20}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)].
\end{aligned}
$$

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

$$
\begin{aligned}
P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{ & e_k' \Lambda_T^{(1)1/2} \tilde{v}_T \ge c_{21}\{\beta_T^{(1)'} \Sigma_{TT}^{(1)-1} \beta_T^{(1)}\}^{-1} \\
& e_k' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \beta_T^{(1)} - c_{22}[\log\{(p_n - q_n) s_n \log n\}/(n\lambda_n^2)]^{1/2} \{\log(q_n s_n \log n)/n\}^{1/2} \\
& - c_{22}[\log\{(p_n - q_n) s_n \log n\}/(n\lambda_n)]\, |e_k' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| \\
& - c_{22}[\log\{(p_n - q_n) s_n \log n\}/(n\lambda_n)]\, [\{q_n s_n \log(q_n s_n)/n\}^{1/2} + \{q_n s_n \log\log(n)/n\}^{1/2}] \Big\}\Big] \\
\ge\ & 1 - c_{23}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)].
\end{aligned}
$$

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

$$
P\Big[\bigcap_{k\in\mathcal{S}_1}\Big\{ e_k' \Lambda_T^{(1)1/2} \tilde{v}_T > 0 \Big\}\Big] \ge 1 - c_{23}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)].
$$

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

$$
P\Big[\bigcap_{k\in\mathcal{S}_2}\Big\{ e_k' \Lambda_T^{(1)1/2} \tilde{v}_T < 0 \Big\}\Big] \ge 1 - c_{24}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)].
$$

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

$$P\{\mathrm{sgn}(\tilde{v}_T) = \mathrm{sgn}(\beta_T^{(1)}) = \mathrm{sgn}(\hat{\beta}_T^{(1)})\}$$

$$\geq 1 - c_{25}[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n^{\kappa_1}/12) + \exp(-n^{\kappa_2}/12)],$$

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

$$\lambda_{\max}(A) \leq \lambda_{\min}(A - B) + \lambda_{\max}(B) \leq \lambda_{\max}(B);$$

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

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 $T = \{1, \dots, q_n\}$. Assume the matrix $\Sigma^{(1)}_{TT}$ is invertible and consider the following optimization problem:*

$$\min_{w_T \in \mathbb{R}^{q_n s_n}} \Big[\frac{1}{2} w_T' \big(\Sigma^{(1)}_{TT} + {}_{1}{}_{2}{}^{(1)}_{T}{}^{(1)'}_{T} \big) w_T - {}_{1}{}_{2} w_T' {}^{(1)}_{T} + {}_{n}(\Lambda^{(1)1/2}_{T} w_T)' sgn({}^{(1)}_{T}) \Big];$$

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

$$\tilde{w}_T = {}_{1}{}_{2}(1 + {}_{n}\|\Lambda^{(1)1/2}_{T}{}^{(1)}_{T}\|_1)(1 + {}_{1}{}_{2}{}^{(1)'}_{T}\Sigma^{(1)}_{TT}{}^{(1)}_{T})^{-1}{}^{(1)}_{T} - {}_{n}\Sigma^{(1)-1}_{TT}\Lambda^{(1)1/2}_{T} sgn({}^{(1)}_{T}).$$

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

$$\big(\Sigma^{(1)}_{TT} + {}_{1}{}_{2}{}^{(1)}_{T}{}^{(1)'}_{T}\big)\tilde{w}_T = {}_{1}{}_{2}{}^{(1)}_{T} - {}_{n}\Lambda^{(1)1/2}_{T}\mathrm{sgn}({}^{(1)}_{T}). \tag{94}$$

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

$$\begin{aligned}\big(\Sigma^{(1)}_{TT} + {}_{1}{}_{2}{}^{(1)}_{T}{}^{(1)'}_{T}\big)^{-1} &= \Sigma^{(1)-1}_{TT} - \Sigma^{(1)-1}_{TT}{}^{(1)}_{T}({}_{1}^{-1}{}_{2}^{-1} + {}^{(1)'}_{T}\Sigma^{(1)-1}_{TT}{}^{(1)}_{T})^{-1}{}^{(1)'}_{T}\Sigma^{(1)-1}_{TT} \\ &= \Sigma^{(1)-1}_{TT} - {}_{1}{}_{2}(1 + {}_{1}{}_{2}{}^{(1)'}_{T}\Sigma^{(1)-1}_{TT}{}^{(1)}_{T})^{-1}\Sigma^{(1)-1}_{TT}{}^{(1)}_{T}{}^{(1)'}_{T}\Sigma^{(1)-1}_{TT}.\end{aligned}$$

Finally, by combining the above two equations, we have

$$
\begin{aligned}
\tilde{w}_T = {} & \Sigma^{(1)}_{TT} + {}_1 {}_2 {}^{(1)}_T {}^{(1)\prime}_T {}^{-1} \{ {}_1 {}_2 {}^{(1)}_T \quad {}_n\Lambda^{(1)1/2}_T \mathrm{sgn}({}^{(1)}_T)\} \\
= {} & \{\Sigma^{(1)-1}_{TT} \quad {}_1 {}_2 (1 + {}_1 {}_2 {}^{(1)\prime}_T \Sigma^{(1)-1}_{TT} {}^{(1)}_T)^{-1} \Sigma^{(1)-1}_{TT} {}^{(1)}_T {}^{(1)\prime}_T \Sigma^{(1)-1}_{TT}\} \\
& \{ {}_1 {}_2 {}^{(1)}_T \quad {}_n\Lambda^{(1)1/2}_T \mathrm{sgn}({}^{(1)}_T)\} \\
= {} & \{\Sigma^{(1)-1}_{TT} \quad {}_1 {}_2 (1 + {}_1 {}_2 {}^{(1)\prime}_T \Sigma^{(1)}_{TT} {}^{(1)}_T)^{-1} {}^{(1)}_T {}^{(1)\prime}_T \Sigma^{(1)-1}_{TT}\} \\
& \{ {}_1 {}_2 {}^{(1)}_T \quad {}_n\Lambda^{(1)1/2}_T \mathrm{sgn}({}^{(1)}_T)\} \\
= {} & \{ {}_1 {}_2 \Sigma^{(1)-1}_{TT} {}^{(1)}_T \quad {}^2_1 {}^2_2 (1 + {}_1 {}_2 {}^{(1)\prime}_T \Sigma^{(1)}_{TT} {}^{(1)}_T)^{-1} {}^{(1)}_T {}^{(1)\prime}_T \Sigma^{(1)-1}_{TT} {}^{(1)}_T\} \\
& \{ {}_n \Sigma^{(1)-1}_{TT} \Lambda^{(1)1/2}_T \mathrm{sgn}({}^{(1)}_T) \quad {}_1 {}_2 (1 + {}_1 {}_2 {}^{(1)\prime}_T \Sigma^{(1)}_{TT} {}^{(1)}_T)^{-1} {}^{(1)}_T \; {}_n \|\Lambda^{(1)1/2}_T {}^{(1)}_T\|_1\} \\
= {} & {}_1 {}_2 (1 + {}_n \|\Lambda^{(1)1/2}_T {}^{(1)}_T\|_1)(1 + {}_1 {}_2 {}^{(1)\prime}_T \Sigma^{(1)}_{TT} {}^{(1)}_T)^{-1} {}^{(1)}_T \quad {}_n \Sigma^{(1)-1}_{TT} \Lambda^{(1)1/2}_T \mathrm{sgn}({}^{(1)}_T),
\end{aligned}
$$

which finishes the proof. □

Lemma 17. *Consider the following optimization problem:*

$$
\min_{w \in \mathbb{R}^{p_n s_n}} \Big[\frac{1}{2} w' \; \Sigma^{(1)} + {}_1 {}_2 {}^{(1)} {}^{(1)\prime} \; w \quad {}_1 {}_2 w' {}^{(1)} + {}_n \sum_{j=1}^{p_n} \|\Lambda^{(1)1/2}_j w_j\|_1 \Big], \tag{95}
$$

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) $\Sigma^{(1)}_{TT}$ *is invertible.*

(b) ${}_1 {}_2 (1 + {}_n \|\Lambda^{(1)1/2}_T {}^{(1)}_T\|_1)(1 + {}_1 {}_2 {}^{(1)\prime}_T \Sigma^{(1)}_{TT} {}^{(1)}_T)^{-1} (\min_{j \in T} \min_{k \le s_n} \omega^{1/2}_{jk} | {}_{jk}|) >$ ${}_n \|\Lambda^{(1)1/2}_T \Sigma^{(1)-1}_{TT} \Lambda^{(1)1/2}_T \mathrm{sgn}({}^{(1)}_T)\|_\infty$.

(c) $\|\Lambda^{(1)-1/2}_N \Sigma^{(1)}_{NT} \Sigma^{(1)-1}_{TT} \Lambda^{(1)1/2}_T \mathrm{sgn}({}^{(1)}_T)\|_\infty \quad 1$ *, for a universal constant* $\in (0, 1]$.

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*$(\ ^{(1)})$.

Proof of Lemma 17: First of all, based on (a), (b) and the definition of $\hat{w}$, it is trivial to deduce that $\mathrm{sgn}(\hat{w}) = \mathrm{sgn}(\ ^{(1)})$, 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

$$\Sigma_{TT}^{(1)} + \ _1\ _2\ _T^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T = \ _1\ _2\ _T^{(1)} \quad _n\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)}); \tag{96}$$

$$k\Lambda_N^{(1)-1/2} \quad \Sigma_{NT}^{(1)} + \ _1\ _2\ _N^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T \quad _1\ _2\ _N^{(1)}\ k_\infty \quad _n; \tag{97}$$

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

$$\begin{aligned}
&k\Lambda_N^{(1)-1/2} \quad \Sigma_{NT}^{(1)} + \ _1\ _2\ _N^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T \quad _1\ _2\ _N^{(1)}\ k_\infty \\
=&k\Lambda_N^{(1)-1/2} \quad \Sigma_{NT}^{(1)} + \ _1\ _2\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\ _T^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T \quad _1\ _2\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\ _T^{(1)}\ k_\infty \\
=&k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)} \quad I + \ _1\ _2\ _T^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T \quad _1\ _2\ _T^{(1)}\ k_\infty;
\end{aligned}$$

where the first and the second equalities follow from (10) in the main paper. For the term $I + \ _1\ _2\ _T^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T$, we have

$$\begin{aligned}
&I + \ _1\ _2\ _T^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T \\
=&\ _1\ _2(1 + \ _nk\Lambda_T^{(1)1/2}\ _T^{(1)}k_1)\ _T^{(1)} \quad _n\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)}) \\
&\quad _1\ _2\ _n\ _T^{(1)}\ _T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)}) \\
=&\ _1\ _2\ _T^{(1)} \quad _n\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)});
\end{aligned}$$

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

$$\begin{aligned}
&k\Lambda_N^{(1)-1/2} \quad \Sigma_{NT}^{(1)} + \ _1\ _2\ _N^{(1)}\ _T^{(1)\prime}\ \tilde{w}_T \quad _1\ _2\ _N^{(1)}\ k_\infty \\
=&\ _nk\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\ _T^{(1)})k_\infty \quad _n;
\end{aligned}$$

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. □

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

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

Assume the condition (a):

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

Then we have the following property:

$$P\{\mathcal{M}_{1n}(\epsilon)\} \geq 1 - 2\epsilon - 2\exp(-n\delta_1/12) - 2\exp(-n\delta_2/12), \quad \forall \epsilon \in (e^{-n/100}, 1/100).$$

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

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

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 $\hat{\beta}_T^{(1)} \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$ is independent of $(n-2)S_{TT}^{(1)} \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$. Together with (98), condition (a) and Theorem 3.2.12 in Muirhead (1982), we reach a conclusion that

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

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

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

For any $\% \, 2 \, (e^{-n/100}; 1{=}100)$, we plug $t = f16(n \quad q_n s_n \quad 1)^{-1} \log(\%^{-1}){=}3g^{1/2}$ into the above inequality to obtain

$$
\begin{aligned}
P \; & j(n \quad q_n s_n \quad 1)^{-1}(n \quad 2)(\hat{}^{(1)\prime}_{T} \Sigma^{(1)-1}_{TT} \hat{}^{(1)}_{T})(\hat{}^{(1)\prime}_{T} S^{(1)-1}_{TT} \hat{}^{(1)}_{T})^{-1} \quad 1j \\
& f16(n \quad q_n s_n \quad 1)^{-1} \log(\%^{-1}){=}3g^{1/2} j f Y_i = y_i g^{n}_{i=1} \setminus M_n \quad 2\%;
\end{aligned}
$$

which implies that

$$
\begin{aligned}
P \; & j(n \quad q_n s_n \quad 1)^{-1}(n \quad 2)(\hat{}^{(1)\prime}_{T} \Sigma^{(1)-1}_{TT} \hat{}^{(1)}_{T})(\hat{}^{(1)\prime}_{T} S^{(1)-1}_{TT} \hat{}^{(1)}_{T})^{-1} \quad 1j \\
& f16(n \quad q_n s_n \quad 1)^{-1} \log(\%^{-1}){=}3g^{1/2} j f Y_i = y_i g^{n}_{i=1} \setminus M_n \quad 1 \quad 2\%:
\end{aligned} \tag{99}
$$

Therefore, it can be seen that

$$
\begin{aligned}
P \; & j(n \quad q_n s_n \quad 1)^{-1}(n \quad 2)(\hat{}^{(1)\prime}_{T} \Sigma^{(1)-1}_{TT} \hat{}^{(1)}_{T})(\hat{}^{(1)\prime}_{T} S^{(1)-1}_{TT} \hat{}^{(1)}_{T})^{-1} \quad 1j \\
& f16(n \quad q_n s_n \quad 1)^{-1} \log(\%^{-1}){=}3g^{1/2} \\
& \sum_{\{y_i\}_{i=1}^{n} \in \mathcal{M}_n} P \; j(n \quad q_n s_n \quad 1)^{-1}(n \quad 2)(\hat{}^{(1)\prime}_{T} \Sigma^{(1)-1}_{TT} \hat{}^{(1)}_{T})(\hat{}^{(1)\prime}_{T} S^{(1)-1}_{TT} \hat{}^{(1)}_{T})^{-1} \\
& \quad 1j \quad f16(n \quad q_n s_n \quad 1)^{-1} \log(\%^{-1}){=}3g^{1/2} j f Y_i = y_i g^{n}_{i=1} \quad P \; f Y_i = y_i g^{n}_{i=1} \\
& (1 \quad 2\%) \sum_{\{y_i\}_{i=1}^{n} \in \mathcal{M}_n} P \; f Y_i = y_i g^{n}_{i=1} = (1 \quad 2\%) P(M_n) \\
& 1 \quad 2\% \quad 2\exp(\; n \; _1{=}12) \quad 2\exp(\; n \; _2{=}12);
\end{aligned} \tag{100}
$$

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 $\% \, 2 \, (e^{-n/100}; 1{=}100)$,

$$
M^{*}_{1n}(\%) \quad M_{1n}(\%); \tag{101}
$$

in which $M^{*}_{1n}(\%) = \; j(n \quad q_n s_n \quad 1)^{-1}(n \quad 2)(\hat{}^{(1)\prime}_{T} \Sigma^{(1)-1}_{TT} \hat{}^{(1)}_{T})(\hat{}^{(1)\prime}_{T} S^{(1)-1}_{TT} \hat{}^{(1)}_{T})^{-1} \quad 1j$ $f16(n \quad q_n s_n \quad 1)^{-1} \log(\%^{-1}){=}3g^{1/2}$. Finally, the assertion follows immediately from (100) and (101). □

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

$$\begin{aligned}
&(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) + 2(n_1^{-1}n_2^{-1}n^2)\{\log(\delta^{-1})/n\} \\
&+ 2(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n + 2n_1 n_2 n^{-2}\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)})^{1/2}\{\log(\delta^{-1})/n\}^{1/2} \\
&\le (400\lambda_1^{-1}\lambda_2^{-1})^{1/2}\left[q_n s_n/n + \log(\delta^{-1})/n + \{\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\log(\delta^{-1})/n\}^{1/2}\right],
\end{aligned}$$

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

$$\begin{aligned}
&P\Big\{\hat\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\hat\beta_T^{(1)} - \beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)} \ge (400\lambda_1^{-1}\lambda_2^{-1})^{1/2}\big[q_n s_n/n + \log(\delta^{-1})/n \\
&+ \{\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\log(\delta^{-1})/n\}^{1/2}\big] \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big\} \le 1 - \delta. \qquad (103)
\end{aligned}$$

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

$$\begin{aligned}
&P\Big\{\hat\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\hat\beta_T^{(1)} - \beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)} \le (n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) - 2(n_1^{-1}n_2^{-1}n^2) \\
&(q_n s_n/n + 2n_1 n_2 n^{-2}\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)})^{1/2}(t/n)^{1/2} \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big\} \\
&\quad \le \exp(-t).
\end{aligned}$$

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

$$\begin{aligned}
&P\Big[\hat\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\hat\beta_T^{(1)} - \beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)} \ge (n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) - 2(n_1^{-1}n_2^{-1}n^2) \\
&(q_n s_n/n + 2n_1 n_2 n^{-2}\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)})^{1/2}\{\log(\delta^{-1})/n\}^{1/2} \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\Big] \ge 1 - \delta.
\end{aligned}$$

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

$$\begin{aligned}
&(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n) - 2(n_1^{-1}n_2^{-1}n^2)(q_n s_n/n + 2n_1 n_2 n^{-2}\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)})^{1/2}\{\log(\delta^{-1})/n\}^{1/2} \\
&\ge -(400\lambda_1^{-1}\lambda_2^{-1})^{1/2}\left[\log(\delta^{-1})/n + \{\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\log(\delta^{-1})/n\}^{1/2}\right].
\end{aligned}$$

We then derive from the above two inequalities that

$$P\Big\{\hat{\theta}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\hat{\theta}_T^{(1)} \geq \theta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\theta_T^{(1)} - (400\lambda_1^{-1}\lambda_2^{-1})^{1/2}\{\log(\delta^{-1})/n$$
$$+\{\theta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\theta_T^{(1)}\log(\delta^{-1})/n\}^{1/2} \mid \{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\Big\} \geq 1-\delta.$$

Together with (103), we arrive at

$$P\left(\mathcal{M}_{2n}(\delta) \mid \{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\right) \geq 1-2\delta. \tag{104}$$

Finally, we have

$$\begin{aligned}
P\{\mathcal{M}_{2n}(\delta)\} &\geq P\{\mathcal{M}_{2n}(\delta)\cap\mathcal{M}_n\} = \sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n} P\left(\mathcal{M}_{2n}(\delta) \mid \{Y_i=y_i\}_{i=1}^n\right) P\left(\{Y_i=y_i\}_{i=1}^n\right)\\
&\geq (1-2\delta)\sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n} P\left(\{Y_i=y_i\}_{i=1}^n\right) = (1-2\delta)P(\mathcal{M}_n)\\
&\geq 1-2\delta-2\exp(-n\lambda_1/12)-2\exp(-n\lambda_2/12),
\end{aligned}$$

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

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

$$\mathcal{M}_{4n}(\delta) = \bigcap_{j=1}^{q_ns_n}\Big\{\left|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\theta}_T^{(1)} - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\theta_T^{(1)}\right| \leq (8\lambda_1^{-1}\lambda_2^{-1})^{1/2}$$
$$\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2}\{\log(q_ns_n\delta^{-1})/n\}^{1/2}\Big\},$$

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

$$P\{\mathcal{M}_{4n}(\delta)\} \geq 1-2\delta-2\exp(-n\lambda_1/12)-2\exp(-n\lambda_2/12), \quad \forall\delta\in(e^{-n/100}, 1/100).$$

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

$\mathcal{M}_n$

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

Moreover, it can be observed that

$$P\left(\mathcal{M}_{4n}(\varrho) \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) \geq \sum_{j=1}^{q_n s_n} P\left(\mathcal{M}_{4nj}(\varrho) \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) - (q_n s_n - 1),$$

where the events $\mathcal{M}_{4nj}(\varrho) = \Big\{\big|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat\beta_T^{(1)} - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\beta_T^{(1)}\big| \leq (8\kappa_1^{-1}\kappa_2^{-1})^{1/2}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2}\{\log(q_n s_n\varrho^{-1})/n\}^{1/2}\Big\}$ for all $j \leq q_n s_n$. Under (105), the concentration inequality entails that for all $j \leq q_n s_n$

$$P\left(\mathcal{M}_{4nj}(\varrho) \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) \geq 1 - 2\exp\{-\log(q_n s_n \varrho^{-1})\} = 1 - 2q_n^{-1}s_n^{-1}\varrho.$$

Putting the above two inequalities together leads to

$$P\left(\mathcal{M}_{4n}(\varrho) \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) \geq 1 - 2\varrho. \tag{106}$$

Therefore, we have

$$\begin{aligned}
P\{\mathcal{M}_{4n}(\varrho)\} &\geq P\{\mathcal{M}_{4n}(\varrho) \cap \mathcal{M}_n\} \\
&\geq (1 - 2\varrho)\sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\left(\{Y_i = y_i\}_{i=1}^n\right) = (1 - 2\varrho)P(\mathcal{M}_n) \\
&\geq 1 - 2\varrho - 2\exp(-n\kappa_1/12) - 2\exp(-n\kappa_2/12),
\end{aligned}$$

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

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

$$\begin{aligned}
\mathcal{M}_{5n}(\varrho) = \Big\{&\big|\hat\beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) - \beta_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\big| \\
&\leq (8\kappa_1^{-1}\kappa_2^{-1})^{1/2}\lambda_{\max}^{1/2}(\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\{q_n s_n \log(\varrho^{-1})/n\}^{1/2}\Big\}.
\end{aligned}$$

Then we have the following property:

$$P\{\mathcal{M}_{5n}(\epsilon)\} \geq 1 - 2\epsilon - 2\exp(-n\lambda_1/12) - 2\exp(-n\lambda_2/12), \quad \forall \epsilon \in (e^{-n/100}, 1/100).$$

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

$$\hat{\nu}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}) \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n$$
$$\sim N\left(\nu_T^{(1)\prime}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)}), n_1^{-1}n_2^{-1}n\{\mathrm{sgn}(\beta_T^{(1)})^\prime\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\}\right).$$

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

$$P\left(\left|(\hat{\nu}_T^{(1)} - \nu_T^{(1)})^\prime\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}\mathrm{sgn}(\beta_T^{(1)})\right| \leq t \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right)$$
$$\geq 1 - 2\exp\left[-8^{-1}\lambda_1\lambda_2\{q_n s_n \lambda_{\max}(\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\}^{-1}nt^2\right].$$

Plugging $t = (8\lambda_1^{-1}\lambda_2^{-1})^{1/2}\lambda_{\max}^{1/2}(\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2})\{q_n s_n \log(\epsilon^{-1})/n\}^{1/2}$ into the above inequality yields

$$P\left(\mathcal{M}_{5n}(\epsilon) \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n\right) \geq 1 - 2\epsilon. \tag{107}$$

Finally, we have

$$\begin{aligned} P\{\mathcal{M}_{5n}(\epsilon)\} &\geq P\{\mathcal{M}_{5n}(\epsilon) \cap \mathcal{M}_n\} \\ &\geq (1 - 2\epsilon)\sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P\left(\{Y_i = y_i\}_{i=1}^n\right) = (1 - 2\epsilon)P(\mathcal{M}_n) \\ &\geq 1 - 2\epsilon - 2\exp(-n\lambda_1/12) - 2\exp(-n\lambda_2/12), \end{aligned}$$

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) $P\left(\max_{j\le q_n s_n} (e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j) \quad 1 \right.$

$\left. c_1 \; q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2} \quad 1 \; c_2 \; \{\log(n)\}^{-1} + \exp(\; n \; _1/12) + \exp(\; n \; _2/12) \right.$.

2) $P\left(\max_{j\le q_n s_n} (e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j) \quad 1 \right.$

$\left. c_1 \; q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2} \quad 1 \; c_2 \; \{\log(n)\}^{-1} + \exp(\; n \; _1/12) + \exp(\; n \; _2/12) \right.$.

3) $P\left(\{sgn(\; _T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\; _T^{(1)})\}/\{sgn(\; _T^{(1)})'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\; _T^{(1)})\} \right.$

$1 \quad c_1 \; q_n s_n/n + \{\log\log(n)/n\}^{1/2} \quad 1 \quad c_2 \; \{\log(n)\}^{-1} + \exp(\; n \; _1/12) +$

$\left. \exp(\; n \; _2/12) \right.$.

4) $P\left(\{sgn(\; _T^{(1)})'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\; _T^{(1)})\}/\{sgn(\; _T^{(1)})'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\; _T^{(1)})\} \right.$

$1 \quad c_1 \; q_n s_n/n + \{\log\log(n)/n\}^{1/2} \quad 1 \quad c_2 \; \{\log(n)\}^{-1} + \exp(\; n \; _1/12) +$

$\left. \exp(\; n \; _2/12) \right.$.

Recall that $\{e_j : j \quad q_n s_n\}$ *denotes the standard basis for* $\mathbb{R}^{q_n s_n}$.

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 $\{Y_i = y_i\}_{i=1}^n \cap M_n$, and for every $j \quad q_n s_n$,

$$(n \quad 2)(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}|\{Y_i = y_i\}_{i=1}^n \cap M_n \quad {}^2_{n-q_n s_n-1}.$$

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 $j \quad q_n s_n$,

$$P\left(|(n \quad q_n s_n \quad 1)^{-1}(n \quad 2)(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1} \right.$$

$$\left. 1| \quad t|\{Y_i = y_i\}_{i=1}^n \cap M_n \right) \quad 2\exp\{ \; 3(n \quad q_n s_n \quad 1)t^2/16\};$$

which together with condition (a) implies that

$$
\begin{aligned}
P\Big(&\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}-1\big| \\
&\le 4q_ns_n/n+2t \,\big|\, \{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\Big)\ge 1-2\exp\{-3(n-q_ns_n-1)t^2/16\} \\
&\ge 1-2\exp(-nt^2/16).
\end{aligned}
$$

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

$$
\begin{aligned}
P\Big(&\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}-1\big| \\
&\le 4q_ns_n/n+2t \,\big|\, \{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\Big)\ge 1-2q_ns_n\exp(-nt^2/16).
\end{aligned}
$$

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

$$
\begin{aligned}
P\Big(&\max_{j\le q_ns_n}\big|(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)^{-1}-1\big| \\
&\le 4q_ns_n/n+8\{\log(q_ns_n\log n)/n\}^{1/2} \,\big|\, \{Y_i=y_i\}_{i=1}^n\cap\mathcal{M}_n\Big) \\
&\ge 1-2\{\log(n)\}^{-1}.
\end{aligned}
\tag{108}
$$

It then follows that

$$
\begin{aligned}
P\Big(&\max_{j\le q_ns_n}(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)\le 1 \\
&+8\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\Big) \\
&\ge\sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n}P\Big(\max_{j\le q_ns_n}(e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j)/(e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j) \\
&\le 1+8\big[q_ns_n/n+\{\log(q_ns_n\log n)/n\}^{1/2}\big]\,\big|\,\{Y_i=y_i\}_{i=1}^n\Big)P\big(\{Y_i=y_i\}_{i=1}^n\big) \\
&\ge[1-2\{\log(n)\}^{-1}]\sum_{\{y_i\}_{i=1}^n\in\mathcal{M}_n}P\big(\{Y_i=y_i\}_{i=1}^n\big)=[1-2\{\log(n)\}^{-1}]P(\mathcal{M}_n) \\
&\ge 1-2\big[\{\log(n)\}^{-1}+\exp(-n\kappa_1/12)+\exp(-n\kappa_2/12)\big];
\end{aligned}
$$

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 $\left\{ \max_{j \le q_n s_n} (e'_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) = (e'_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1 \quad 8 \; q_n s_n = n + f \log(q_n s_n \log n) = n g^{1/2} \right\}$, it is straightforward

to verify that

$$\max_{j \le q_n s_n} (e'_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) = (e'_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1$$

$(1)-1$ TT (1) ((1)) n $q \log(ne)$ 277 9701 10 8210 () 217 9701 6 5860 (1) 247 9701 17 407 9 381 (

$n s_n$ Tf 6.988 -292.2d [Td [(Puttingr)501ommaa9 1 [(g0 rg

$$2 \max_{j \le q_n s_n} (e'_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) = (e'_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1 \, :$$

f log

$(1)-1$ TT (1)

Putting the above t[(e)]3 T51251[inequalities][(orc27((the) 3bredsthe) 32o23 11.9552 Tf 3.985 61 1321 Td

log(ne)0 1 Td [( )]whicT512hitcompletest27(ab)-3(pro)27(v)fabpropereyab2TJ/F! Similarabreason where

$$\Omega_{1j} = |e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\xi}_T^{(1)} - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\xi_T^{(1)}|;$$

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

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

$$P\left[\bigcap_{j=1}^{q_n s_n}\left\{\Omega_{1j} \le c_1\{\log(q_n s_n \log n)/n\}^{1/2}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2}\right\}\right]$$
$$\ge 1 - c_2\left[\{\log(n)\}^{-1} + \exp(-n\kappa_1/12) + \exp(-n\kappa_2/12)\right]. \quad (110)$$

Regarding the term Ω_{2j}, it can be seen that

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

where

$$\Pi_{1j} = \{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\xi}_T^{(1)}\}\{e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1}$$
$$- \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\xi}_T^{(1)}\}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1};$$

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

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

$$P\left(\max_{j \le q_n s_n}\Pi_{2j} \le c_3[q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]\right)$$
$$\ge 1 - c_4\left[\{\log(n)\}^{-1} + \exp(-n\kappa_1/12) + \exp(-n\kappa_2/12)\right].$$

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

$$P\Big[\bigcap_{j=1}^{q_n s_n}\Big\{|\Omega_{2j}| \le c_5 \, |\Pi_{1j}| \, \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\} + [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]$$
$$\{\log(q_n s_n \log n)/n\}^{1/2} \, \{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{1/2} + [q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]$$
$$|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\hat{\gamma}_T^{(1)}|\Big\}\Big] \ge 1 - c_6 \, \{\log(n)\}^{-1} + \exp(-n^{\delta_1}/12) + \exp(-n^{\delta_2}/12) \, . \tag{111}$$

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

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

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

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

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

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

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 $j \le q_n s_n$,

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

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

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 $t_j \ge 0$,

$$
\begin{aligned}
&P\left[|\Pi_{1j}| \ge t_j \mid \{Y_i = y_i\}_{i=1}^n \cap \mathcal{M}_n \cap \{\hat{\nu}_T\}\right] \le c_7 \exp\{-c_8(n - q_n s_n - 3)\sigma_j^{-1} t_j^2\} \\
&\le c_7 \exp\left[-2^{-1} c_8 n \{\hat{\nu}_T' \Sigma_{TT}^{(1)-1} \hat{\nu}_T\}^{-1} \{e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j\} t_j^2\right];
\end{aligned}
$$

which further implies that

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

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

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

It then follows that

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

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,

$$P\left[\bigcap_{j=1}^{q_n s_n}\left\{|\Pi_{1j}| \le c_{11}\{e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}e_j\}^{-1/2}\{\log(q_n s_n \log n)/n\}^{1/2}\right.\right.$$

$$\bigcap_{T}^{s_n} q_n s_n/n + \log\log(n)/n + [1 + q_n s_n/n + \{\log\log(n)/n\}^{1/2}]\{{}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}{}_T^{(1)}\}$$

$$+ \{\log\log(n)/n\}^{1/2}\{{}_T^{(1)\prime}\Sigma_{TT}^{(1)-1}{}_T^{(1)}\}^{1/2}\ {}^{1/2}\Big\}\Big] \ \ n + \{\log\log(q_n s_n)(\log n)/n\}^{1/2}\{{}_0\Lambda_T^{(1)1/2}\Sigma^{(1)}$$

$$1 - c_{12}\ (n)\}^{-1} + \exp(-n\ {}_1/12) + \exp(-n\ {}_2/12)\ .$$

Together with (11), it is clear that there exist universal constants

Together with (

(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:*

$$
\begin{aligned}
P\Big[\bigcap_{j=1}^{q_n s_n}\Big\{ & |e_j'\Lambda_T^{(1)1/2}S_{TT}^{(1)-1}\hat{\Lambda}_T^{(1)1/2}sgn(\beta_T^{(1)}) - e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})| \\
& \le c_3\big(\{q_n s_n \log(q_n s_n)/n\}^{1/2} + \{q_n s_n \log\log(n)/n\}^{1/2} \\
& + c_3|e_j'\Lambda_T^{(1)1/2}\Sigma_{TT}^{(1)-1}\Lambda_T^{(1)1/2}sgn(\beta_T^{(1)})| \cdot \sqrt{q_n s_n/n} + \{\log(q_n s_n \log n)/n\}^{1/2}\big)\Big\}\Big] \\
& \ge 1 - c_4[(q_n s_n)^{-1} + \{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)].
\end{aligned}
$$

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

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

where

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

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

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

where c_1 is defined in condition (b), and

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

To bound the term Π_{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 $1 - c_3[\{\log(n)\}^{-1} +$

$\exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)]$,

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

To bound the term Π_{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 $t \ge 0$,

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

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

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

Together with Lemma 3, there exist universal constants $c_8 > 0$ and $c_9 > 0$ such that with probability at least $1 - c_8[\{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)]$,

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

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

$$\begin{aligned} P\Big[\cap_{j=1}^{q_n s_n} \Big\{&\Omega_{1j} \le c_{10}\{q_n s_n \log(q_n s_n \log n)/n\}^{1/2} + c_{10}|e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \mathrm{sgn}(\beta_T^{(1)})| \\ &[q_n s_n/n + \{\log(q_n s_n \log n)/n\}^{1/2}]\Big\}\Big] \\ \ge\ & 1 - c_{11}[\{\log(n)\}^{-1} + \exp(-n\epsilon_1/12) + \exp(-n\epsilon_2/12)]. \end{aligned} \tag{119}$$

To bound the term Ω_{2j}, it can be verified that

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

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

$$P\left(\max_{j \le q_n s_n} \Omega_{2j} \quad c_{12} \{q_n s_n \log(q_n s_n)/n\}^{1/2} \right)$$
$$1 \quad c_{13}[(q_n s_n)^{-1} + \exp(\;\; n \;\; _1/12) + \exp(\;\; n \;\; _2/12)].$$

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

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.