ASYMPTOTIC ANALYSIS OF MIS-CLASSIFIED LINEAR MIXED MODELS

Haiqiang Ma and Jiming Jiang

Jiangxi University of Finance and Economics

and University of California, Davis

Supplementary Material

Throughout this supplementary material, the paper by Ma and Jiang (2022), Asymptotic Analysis of Mis-classified Linear Mixed Models, is referred to as MJ22. All of the notations used below are consistent with those introduced in MJ22.

S1 Some lemmas

The proofs of Theorem 1 and Theorem 2 require the following lemmas, whose proofs are given in the next subsection.

Lemma 1. Let $S_{il} = \sum_{j=1}^{n} 1_{(\gamma_{ij}=l)}$. If $p = O(1/\sqrt{mn})$ and $m/n = O(1)$, then we have

$$\mathrm{Var}\big[\mathrm{tr}\big\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\big\}\big] = \mathrm{Var}\left(\sum_{i=1}^{m}\sum_{l=1}^{m} S_{il}^2\right) = m\mathrm{Var}\left(\sum_{l=1}^{m} S_{1l}^2\right) = mn^3pO(1).$$

Lemma 2. Define $b(\theta_0) = (\sigma_0 I_N \ \tau_0 \tilde{Z})'$. Write $P = I_m \otimes (I_n - n^{-1}J_n)$. If $p = O(1/\sqrt{mn})$ and $m \sim n$, then, we have $\mathrm{E}(\mathrm{tr}[\{b(\theta_0)Pb'(\theta_0)\}^2]) = mnO(1)$.

S2 Proofs of lemmas

S2.1 Proof of Lemma 1

First, we have the following expressions: $\mathrm{Var}\left[\mathrm{tr}\left\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\right\}\right] =$

$$\mathrm{Var}\left\{\sum_{i=1}^{m} \mathrm{tr}(\tilde{Z}_i' J_n \tilde{Z}_i)\right\} = \mathrm{Var}\left[\sum_{i=1}^{m}\sum_{l=1}^{m}\left\{\sum_{j=1}^{n} 1_{(\gamma_{ij}=l)}\right\}^2\right].$$

By the assumption on γ_{ij}, it can be seen that $\sum_{l=1}^m S_{il}^2 = \sum_{l=1}^m\{\sum_{j=1}^n 1_{(\gamma_{ij}=l)}\}^2$ are i.i.d. Thus, we have $\mathrm{Var}[\mathrm{tr}\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\}] = m\mathrm{Var}(\sum_{l=1}^m S_{1l}^2)$.

Now let us consider $\mathrm{Var}(\sum_{l=1}^m S_{1l}^2)$. First, by the assumptions, we have

$$\begin{aligned}
&\mathrm{E}\left(\sum_{l=1}^{m} S_{1l}^2\right) = \mathrm{E}\left(S_{11}^2 + \sum_{l\neq 1} S_{1l}^2\right) = \mathrm{E}\left\{\sum_{j=1}^{n} 1_{(\gamma_{1j}=1)} + \sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\} \\
&+ \sum_{l\neq 1}\mathrm{E}\left\{\sum_{j=1}^{n} 1_{(\gamma_{1j}=l)} + \sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=l)}1_{(\gamma_{1j_2}=l)}\right\} \\
&= n(1-p) + n(n-1)(1-p)^2 + (m-1)\left\{\frac{np}{m-1} + \frac{n(n-1)p^2}{(m-1)^2}\right\} \\
&= n^2(1-p)^2 + 2np - np^2 + \frac{n(n-1)p^2}{(m-1)} = n^2(1-p)^2 + npO(1),
\end{aligned}$$

using the fact that $p = O(1/\sqrt{mn})$ and $m \sim n$ imply that $p = O(1/n)$.

Next, we consider the expectation of $(\sum_{l=1}^m S_{1l}^2)^2$. Note that

$$\mathrm{E}\left(\sum_{l=1}^m S_{1l}^2\right)^2 = \mathrm{E}\left(\sum_{l=1}^m S_{1l}^4 + \sum_{l_1\neq l_2} S_{1l_1}^2 S_{1l_2}^2\right)$$

$$= \mathrm{E}(S_{11}^4) + \sum_{l\neq 1}\mathrm{E}(S_{1l}^4) + \sum_{l_1\neq 1}\mathrm{E}(S_{1l_1}^2 S_{11}^2) + \sum_{l_2\neq 1}\mathrm{E}(S_{1l_2}^2 S_{11}^2) + \sum_{l_1\neq l_2\neq 1}\mathrm{E}(S_{1l_1}^2 S_{1l_2}^2).$$

We study each term in the last expression. First, we have

$$\begin{aligned}
\mathrm{E}(S_{11}^4) &= \mathrm{E}\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=1)} + \sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\}^2 \\
&= \mathrm{E}\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=1)}\right\} + \mathrm{E}\left\{\sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\} \\
&\quad +2\mathrm{E}\left\{\sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\} + 2\mathrm{E}\left\{\sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\} \\
&\quad +2\mathrm{E}\left\{\sum_{j_1\neq j_2\neq j_3} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}\right\} \\
&\quad +\mathrm{E}\left\{\sum_{j_1=j_3\neq j_2=j_4} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\} \\
&\quad +\mathrm{E}\left\{\sum_{j_1=j_4\neq j_2=j_3} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\} \\
&\quad +\mathrm{E}\left\{\sum_{j_1\neq j_2\neq j_3\neq j_4} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\} \\
&\quad +\mathrm{E}\left\{\sum_{j_1=j_3\neq j_2\neq j_4} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\} \\
&\quad +\mathrm{E}\left\{\sum_{j_1=j_4\neq j_2\neq j_3} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\} \\
&\quad +\mathrm{E}\left\{\sum_{j_1\neq j_2=j_3\neq j_4} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\}
\end{aligned}$$

$$
\begin{aligned}
&+\mathrm{E}\left\{\sum_{j_1\neq j_2=j_4\neq j_3} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\} \\
= \;& n(1-p)\{1+7(n-1)(1-p)+6(n-1)(n-2)(1-p)^2 \\
&+(n-1)(n-2)(n-3)(1-p)^3\} \\
= \;& n^4(1-p)^4+n^3pO(1),
\end{aligned}
$$

using the fact that $1+7(n-1)(1-p)+6(n-1)(n-2)(1-p)^2+(n-1)(n-2)(n-3)(1-p)^3 = n^3(1-p)^3+n^2pO(1)$. By similar arguments, it can be shown that

$$
\begin{aligned}
&\mathrm{E}\left(\sum_{l\neq 1} S_{1l}^4\right) = (m-1)\left\{\frac{np}{m-1}+7n(n-1)\frac{p^2}{(m-1)^2}\right. \\
&\left.+6n(n-1)(n-2)\frac{p^3}{(m-1)^3}+n(n-1)(n-2)(n-3)\frac{p^4}{(m-1)^4}\right\} \\
&= np+7n(n-1)\frac{p^2}{(m-1)}+6n(n-1)(n-2)\frac{p^3}{(m-1)^2} \\
&+n(n-1)(n-2)(n-3)\frac{p^4}{(m-1)^3} = npO(1), \\
&\mathrm{E}\left(\sum_{l_1\neq 1} S_{1l_1}^2 S_{11}^2\right) = \sum_{l_1\neq 1}\mathrm{E}\left[\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=l_1)}\right\}^2\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=1)}\right\}^2\right] \\
&= (m-1)\mathrm{E}\left[\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=1)}+\sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\}\right. \\
&\left.\times\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=l_1)}+\sum_{j_3\neq j_4} 1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\right] \\
&= (m-1)\mathrm{E}\left[\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=1)}\right\}\left\{\sum_{j=1}^n 1_{(\gamma_{1j}=l_1)}\right\}\right]
\end{aligned}
$$

$$+(m-1)\mathrm{E}\left[\left\{\sum_{j=1}^{n}1_{(\gamma_{1j}=1)}\right\}\left\{\sum_{j_3\neq j_4}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\right]$$
$$+(m-1)\mathrm{E}\left[\left\{\sum_{j=1}^{n}1_{(\gamma_{1j}=l_1)}\right\}\left\{\sum_{j_1\neq j_2}1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\}\right]$$
$$+(m-1)\mathrm{E}\left[\left\{\sum_{j_1\neq j_2}1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\}\left\{\sum_{j_3\neq j_4}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\right].$$

We now obtain further expressions for the terms appear in the latest expression.

By the assumptions, it can be shown that

$$\mathrm{E}\left[\left\{\sum_{j=1}^{n}1_{(\gamma_{1j}=1)}\right\}\left\{\sum_{j=1}^{n}1_{(\gamma_{1j}=l_1)}\right\}\right]=\frac{n(n-1)p(1-p)}{m-1};$$

$$\begin{aligned}
&\mathrm{E}\left[\left\{\sum_{j=1}^{n}1_{(\gamma_{1j}=1)}\right\}\left\{\sum_{j_3\neq j_4}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\right]\\
&=\mathrm{E}\left\{\sum_{j_3\neq j_4}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}+\sum_{j_3\neq j_4}1_{(\gamma_{1j_4}=1)}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\\
&+\mathrm{E}\left\{\sum_{j\neq j_3\neq j_4}1_{(\gamma_{1j}=1)}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\\
&=\frac{n(n-1)(n-2)p^2(1-p)}{(m-1)^2};
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[\left\{\sum_{j=1}^{n}1_{(\gamma_{1j}=l_1)}\right\}\left\{\sum_{j_1\neq j_2}1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\}\right]\\
&=\mathrm{E}\left\{\sum_{j_3\neq j_4}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}+\sum_{j_3\neq j_4}1_{(\gamma_{1j_4}=l_1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\}\\
&+\mathrm{E}\left\{\sum_{j\neq j_3\neq j_4}1_{(\gamma_{1j}=l_1)}1_{(\gamma_{1j_3}=1)}1_{(\gamma_{1j_4}=1)}\right\}\\
=\;&\frac{n(n-1)(n-2)p(1-p)^2}{(m-1)};
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[\left\{\sum_{j_1\neq j_2} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}\right\}\left\{\sum_{j_3\neq j_4} 1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\right]\\
&= \mathrm{E}\left\{\sum_{j_1\neq j_2\neq j_3\neq j_4} 1_{(\gamma_{1j_1}=1)}1_{(\gamma_{1j_2}=1)}1_{(\gamma_{1j_3}=l_1)}1_{(\gamma_{1j_4}=l_1)}\right\}\\
= \quad & \frac{n(n-1)(n-2)(n-3)p^2(1-p)^2}{(m-1)^2}.
\end{aligned}$$

Combining the above results, we obtain the following:

$$\begin{aligned}
&\mathrm{E}\left(\sum_{l_1\neq 1} S_{1l_1}^2 S_{11}^2\right) = n(n-1)p(1-p) + \frac{n(n-1)(n-2)p^2(1-p)}{m-1}\\
&+n(n-1)(n-2)p(1-p)^2 + \frac{n(n-1)(n-2)(n-3)p^2(1-p)^2}{m-1} = n^3pO(1).
\end{aligned}$$

Similarly, the following expression can be obtained:

$$\begin{aligned}
&\mathrm{E}\left(\sum_{l_1\neq l_2\neq 1} S_{1l_1}^2 S_{1l_2}^2\right) = (m-2)\left\{\frac{n(n-1)p^2}{m-1} + 2\frac{n(n-1)(n-2)p^3}{(m-1)^2}\right.\\
&\left.+\frac{n(n-1)(n-2)(n-3)p^4}{(m-1)^3}\right\} = npO(1).
\end{aligned}$$

Combining the above expressions, we obtain that

$$\mathrm{E}\left(\sum_{l=1}^{m} S_{1l}^2\right)^2 = n^4(1-p)^4 + n^3pO(1).$$

Combining the above results, we have

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{l=1}^{m} S_{1l}^2\right) = \mathrm{E}\left(\sum_{l=1}^{m} S_{1l}^2\right)^2 - \left\{\mathrm{E}\left(\sum_{l=1}^{m} S_{1l}^2\right)\right\}^2\\
&= n^4(1-p)^4 + n^3pO(1) - \{n^2(1-p)^2 + npO(1)\}^2 = n^3pO(1).
\end{aligned}$$

S2.2 Proof of Lemma 2

Straightforward calculation shows that

$$
\begin{aligned}
&\mathrm{tr}[\{b(\theta_0)Pb'(\theta_0)\}^2] = \sigma_0^4\mathrm{tr}(P) + 2\sigma_0^2\tau_0^2\mathrm{tr}(\tilde{Z}'P\tilde{Z}) + \tau_0^4\mathrm{tr}\{(\tilde{Z}'P\tilde{Z})^2\} \\
&= m(n-1)\sigma_0^4 + 2\sigma_0^2\tau_0^2\left[mn - \frac{1}{n}\mathrm{tr}\{\tilde{Z}'(I_m\otimes J_n)\tilde{Z}\}\right] \\
&+\tau_0^4\mathrm{tr}\left[\left\{\tilde{Z}'\tilde{Z} - \frac{1}{n}\tilde{Z}'(I_m\otimes J_n)\tilde{Z}\right\}^2\right].
\end{aligned}
$$

Furthermore, it can be shown that

$$
\begin{aligned}
\mathrm{E}\left[\frac{1}{n}\mathrm{tr}\left\{\tilde{Z}'(I_m\otimes J_n)\tilde{Z}\right\}\right] &= mn(1-p)^2 + 2mp - mp^2 + \frac{m(n-1)p^2}{m-1} \\
&= mn(1-p)^2 + mpO(1).
\end{aligned}
$$

The proof is complete by showing below that

$$
\mathrm{E}\left(\mathrm{tr}\left[\left\{\tilde{Z}'\tilde{Z} - \frac{1}{n}\tilde{Z}'(I_m\otimes J_n)\tilde{Z}\right\}^2\right]\right) = mnpO(1). \qquad \text{(S2.1)}
$$

Let us begin with the following expression:

$$
\begin{aligned}
&\mathrm{tr}(\tilde{Z}'\tilde{Z})^2 = \sum_{l=1}^{m}\left[\sum_{i=1}^{m}\left\{\sum_{j=1}^{n}1_{(\gamma_{ij}=l)}\right\}\right]^2 \\
&= \sum_{l=1}^{m}\sum_{i=1}^{m}\left\{\sum_{j=1}^{n}1_{(\gamma_{ij}=l)}\right\}^2 + \sum_{l=1}^{m}\sum_{i_1\neq i_2}\left\{\sum_{j=1}^{n}1_{(\gamma_{i_1j}=l)}\right\}\left\{\sum_{j=1}^{n}1_{(\gamma_{i_2j}=l)}\right\} \\
&\equiv \sum_{i=1}^{m}S_{ii}^2 + \sum_{i\neq l}S_{il}^2 + \sum_{i_1\neq i_2}S_{i_1i_1}S_{i_2i_1} + \sum_{i_1\neq i_2}S_{i_2i_2}S_{i_1i_2} + \sum_{l\neq i_1\neq i_2}S_{i_1l}S_{i_2l},
\end{aligned}
$$

with the S terms defined in obvious ways. It can be shown that

$$\mathrm{E}\left(\sum_{i=1}^{m} S_{ii}^2\right) = m\{n(1-p) + n(n-1)(1-p)^2\},$$
$$\mathrm{E}\left(\sum_{i\neq l} S_{il}^2\right) = m(m-1)\left\{\frac{np}{m-1} + \frac{n(n-1)p^2}{(m-1)^2}\right\},$$
$$\mathrm{E}\left(\sum_{i_1\neq i_2} S_{i_1 i_1} S_{i_2 i_1}\right) = mn^2 p(1-p),$$
$$\mathrm{E}\left(\sum_{l\neq i_1\neq i_2} S_{i_1 l} S_{i_2 l}\right) = \sum_{i_1\neq i_2\neq l} \mathrm{E}(S_{i_1 l})\mathrm{E}(S_{i_2 l}) = m(m-1)(m-2)\left(\frac{np}{m-1}\right)^2.$$

It can then be shown that $\mathrm{E}\{\mathrm{tr}(\tilde{Z}'\tilde{Z})^2\} = mn^2 + mnpO(1)$.

Next, the following expressions can be derived:

$$\begin{aligned}\left\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\right\}^2 &= \left(\sum_{i=1}^{m} \tilde{Z}_i' J_n \tilde{Z}_i\right)^2 \\ &= \sum_{i=1}^{m} (\tilde{Z}_i' J_n \tilde{Z}_i)^2 + \sum_{i_1\neq i_2} (\tilde{Z}_{i_1}' J_n \tilde{Z}_{i_1})(\tilde{Z}_{i_2}' J_n \tilde{Z}_{i_2}).\end{aligned}$$

The following expressions can be derived: $\mathrm{tr}\{\sum_{i=1}^{m} (\tilde{Z}_i' J_n \tilde{Z}_i)^2\} = \sum_{i=1}^{m}(\sum_{l=1}^{m} S_{il}^2)^2$;

$$\mathrm{tr}\left\{\sum_{i_1\neq i_2} (\tilde{Z}_{i_1}' J_n \tilde{Z}_{i_1})(\tilde{Z}_{i_2}' J_n \tilde{Z}_{i_2})\right\} = \sum_{i_1\neq i_2}\left(\sum_{l=1}^{m} S_{i_1 l} S_{i_2 l}\right)^2;$$

$$\begin{aligned}\mathrm{E}[\mathrm{tr}\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\}^2] &= \sum_{i=1}^{m} \mathrm{E}\left(\sum_{l=1}^{m} S_{il}^2\right)^2 + \sum_{i_1\neq i_2} \mathrm{E}\left(\sum_{l=1}^{m} S_{i_1 l} S_{i_2 l}\right)^2 \\ &= m\mathrm{E}\left(\sum_{l=1}^{m} S_{1l}^2\right)^2 + m(m-1)\mathrm{E}\left(\sum_{l=1}^{m} S_{1l} S_{2l}\right)^2.\end{aligned}$$

We now evaluate the expectations in the last expression. According to the proof of Lemma 1, the first expectation is $n^4(1-p)^4 + n^3 pO(1)$. As for the second expectation, we have

$$\begin{aligned}
&\mathrm{E}\left(\sum_{l=1}^{m} S_{1l}S_{2l}\right)^2 \\
&= \sum_{l=1}^{m} \mathrm{E}(S_{1l}^2 S_{2l}^2) + \sum_{l_1 \neq l_2} \mathrm{E}(S_{1l_1}S_{1l_2})\mathrm{E}(S_{2l_1}S_{2l_2}) \\
&= 2\{n(1-p) + n(n-1)(1-p)^2\}\left\{\frac{np}{m-1} + \frac{n(n-1)p^2}{(m-1)^2}\right\} \\
&+(m-2)\left\{\frac{np}{m-1} + \frac{n(n-1)p^2}{(m-1)^2}\right\}^2 + (m-2)(m-3)\left\{\frac{n(n-1)p^2}{(m-1)^2}\right\}^2 \\
&+4\left[\left\{\frac{n(n-1)p(1-p)}{m-1}\right\}^2 + \frac{n^2(n-1)^2p^3(1-p)}{(m-1)^3}\right] = n^2pO(1).
\end{aligned}$$

Thus, under the assumptions of the lemma, we have

$$\begin{aligned}
\mathrm{E}\left(\mathrm{tr}[\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\}^2]\right) &= mn^4(1-p)^4 + mn^3pO(1) \\
&= mn^4 - 4mn^4p + mn^3pO(1).
\end{aligned}$$

Finally, the following expression can be derived:

$$\mathrm{tr}\{\tilde{Z}'\tilde{Z}\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\} = \sum_{i=1}^{m}\sum_{l=1}^{m} S_{il}^3 + \sum_{i_1 \neq i_2}\sum_{l=1}^{m} S_{i_1 l}S_{i_2 l}^2.$$

By similar arguments, one can show $\mathrm{E}\left(\sum_{i_1 \neq i_2}\sum_{l=1}^{m} S_{i_1 l}S_{i_2 l}^2\right) = mn^3p + mn^2pO(1)$,

$$\begin{aligned}
&\mathrm{E}\left(\sum_{i=1}^{m}\sum_{l=1}^{m} S_{il}^3\right) = mn(1-p)\{1 + 3(n-1)(1-p) + (n-1)(n-2)(1-p)^2\} \\
&+mnpO(1) = mn^3(1-p)^3 + mn^2pO(1) = mn^3 - 3mn^3p + mn^2pO(1),
\end{aligned}$$

noting that $1+3(n-1)(1-p)+(n-1)(n-2)(1-p)^2 = n^2(1-p)^2+npO(1)$.

Combining the above results, we see the left side of (S2.1) is equal to

$$\mathrm{E}\left[\mathrm{tr}\{(\tilde{Z}'\tilde{Z})^2\}\right] - \frac{2}{n}\mathrm{E}\left[\mathrm{tr}\{\tilde{Z}'\tilde{Z}\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\}\right] + \frac{1}{n^2}\mathrm{E}\left(\mathrm{tr}[\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\}^2\right)$$
$$= mn^2 + mnpO(1) - \frac{2}{n}\{mn^3 - 2mn^3p + mn^2pO(1)\}$$
$$+\frac{1}{n^2}\{mn^4 - 4mn^4p + mn^3pO(1)\} = mnpO(1).$$

S3 Proof of Theorem 1

We provide the proof for the REML part. We have

$$\begin{aligned} \log f_\theta(\tilde{y}) &= -\frac{N-1}{2}\log(2\pi) - \frac{1}{2}\log|\Phi'(ZZ'\tau^2 + I_N\sigma^2)\Phi| \\ &\quad -\frac{1}{2}\tilde{y}'\{\Phi'(ZZ'\tau^2 + I_N\sigma^2)\Phi\}^{-1}\tilde{y} \\ &= -\frac{N-1}{2}\log(2\pi) - \frac{1}{2}l_\theta(\tilde{y}), \end{aligned} \tag{S3.1}$$

with $l_\theta(y) = \log|\Phi'(ZZ'\tau^2 + I_N\sigma^2)\Phi| + \tilde{y}'\{\Phi'(ZZ'\tau^2 + I_N\sigma^2)\Phi\}^{-1}\tilde{y}$. We need to look for the sequences of positive numbers, $p_l(N)$, $l = 1, 2$ of Lemma 7.2 of Jiang (1996). By Lemma 4.2 in Jiang (1996) and the assumptions of Theorem 1, it can be shown that $p_i(N) \sim \|V_i\|_{\mathrm{R}}, i = 1, 2$, where

$$\|V_i\|_{\mathrm{R}}^2 \sim \mathrm{E}_{\theta_0}\left[\left.\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i^2}\right|_{\theta=\theta_0}\right].$$

Thus, it suffices to obtain the orders of $\|V_1\|_{\mathrm{R}}$ and $\|V_2\|_{\mathrm{R}}$.

By matrix differentiation (e.g., Jiang and Nguyen (2021), sec.A.2), we have

$$\frac{\partial l_\theta(\tilde{y})}{\partial \theta_1} = \frac{\partial l_\theta(\tilde{y})}{\partial \tau^2} = \mathrm{tr}[\{\Phi'(\tau^2 ZZ' + \sigma^2 I_N)\Phi\}^{-1}\Phi' ZZ'\Phi]$$
$$-\tilde{y}'\{\Phi'(\tau^2 ZZ' + \sigma^2 I_N)\Phi\}^{-1}\Phi' ZZ'\Phi\{\Phi'(\tau^2 ZZ' + \sigma^2 I_N)\Phi\}^{-1}\tilde{y},$$
$$\frac{\partial l_\theta(\tilde{y})}{\partial \theta_2} = \frac{\partial l_\theta(\tilde{y})}{\partial \sigma^2} = \mathrm{tr}[\{\Phi'(\tau^2 ZZ' + \sigma^2 I_N)\Phi\}^{-1}(\Phi'\Phi)]$$
$$-\tilde{y}'\{\Phi'(\tau^2 ZZ' + \sigma^2 I_N)\Phi\}^{-1}(\Phi'\Phi)\{\Phi'(\tau^2 ZZ' + \sigma^2 I_N)\Phi\}^{-1}\tilde{y}.$$

Let $\mathcal{V}(\theta) = \tau^2 ZZ' + \sigma^2 I_N$, $\mathcal{P}(\theta) = \Phi\{\Phi'(\tau^2 ZZ' + \sigma^2 I_N)\Phi\}^{-1}\Phi'$, by a matrix identity [e.g., Jiang and Nguyen (2021), eq. (1.11)], we have

$$\mathcal{P}(\theta) = \mathcal{V}^{-1}(\theta) - \mathcal{V}^{-1}(\theta)X'\{X\mathcal{V}^{-1}(\theta)X'\}^{-1}X\mathcal{V}^{-1}(\theta)$$
$$= (ZZ'\tau^2 + I_N\sigma^2)^{-1}$$
$$-(ZZ'\tau^2 + I_N\sigma^2)^{-1}X'\{X(ZZ'\tau^2 + I_N\sigma^2)^{-1}X'\}^{-1}X(ZZ'\tau^2 + I_N\sigma^2)^{-1}.$$

The following expressions can then be derived:

$$\frac{\partial l_\theta(\tilde{y})}{\partial \tau^2} = \mathrm{tr}\{\mathcal{P}(\theta)ZZ'\} - y'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)y,\ \frac{\partial l_\theta(\tilde{y})}{\partial \sigma^2} = \mathrm{tr}\{\mathcal{P}(\theta)\} - y'\mathcal{P}^2(\theta)y.$$

Differentiating again, the following expressions can be derived:

$$\begin{aligned}\frac{\partial^2 l_\theta(\tilde{y})}{\partial \tau^4} &= -\mathrm{tr}\{\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\} + 2y'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)y,\\ \frac{\partial^2 l_\theta(\tilde{y})}{\partial \sigma^4} &= -\mathrm{tr}\{\mathcal{P}^2(\theta)\} + 2y'\mathcal{P}^3(\theta)y. \qquad \text{(S3.2)}\end{aligned}$$

Under the true one-way random effects model (2.4) of MJ22, we have $y - X\mu = \tilde{Z}\alpha + \epsilon$. Define $b(\theta) = (\sigma I_N, \tau\tilde{Z})'$, $W_{Nl}(\theta_0) = \epsilon_l/\sigma_0$, $1 \le l \le N$, and $W_{Nl}(\theta_0) = \alpha_{l-N}/\tau_0$, $N+1 \le l \le N+m$. Then, by the assumptions of Theorem 1, it follows that $W_{Nl}(\theta_0)$'s are distributed independently as $\mathrm{N}(0,1)$, independently with $b(\theta_0)$, and $\tilde{y} = \Phi' b'(\theta_0) W_N(\theta_0)$ with

$$W_N(\theta_0) = [W_{N1}(\theta_0), \cdots, W_{N,N+m}(\theta_0)]'.$$

For $\|V_1\|_{\mathrm{R}}^2$, note that

$$\left.\frac{\partial^2 l_\theta(\tilde{y})}{\partial \theta_1^2}\right|_{\theta=\theta_0} = \left.\frac{\partial^2 l_\theta(\tilde{y})}{\partial \tau^4}\right|_{\theta=\theta_0}$$
$$= -\mathrm{tr}\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\} + 2y'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)y$$
$$\equiv -\mathrm{tr}\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\}$$
$$+2W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0).$$

By properties of the normal quadratic forms, we have

$$\mathrm{E}_{\theta_0}\left[\left.\frac{\partial^2 l_\theta(\tilde{y})}{\partial \theta_2^2}\right|_{\theta=\theta_0}\right]$$
$$= -\mathrm{tr}\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\}$$
$$+2\mathrm{E}_{\theta_0}\left[\mathrm{E}\{W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0)|\tilde{Z}\}\right]$$
$$= -\mathrm{tr}\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\} + 2\mathrm{tr}\left[\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\mathrm{E}\{b'(\theta_0)b(\theta_0)\}\right]$$
$$\equiv \Pi_{11} + \Pi_{12}. \tag{S3.3}$$

Now note that $X = 1_m \otimes 1_n$, $ZZ' = I_m \otimes J_n$, $\mathcal{V}(\theta_0) = \tau_0^2 ZZ' + \sigma_0^2 I_N$, and $\mathcal{V}^{-1}(\theta_0) = \sigma_0^{-2} I_m \otimes \{I_n - \tau_0^2(n\tau_0^2 + \sigma_0^2)^{-1}J_n\}$, implying $X'\mathcal{V}^{-1}(\theta)X =$

$mn(n\tau_0^2+\sigma_0^2)^{-1}$, and

$$\mathcal{P}(\theta_0)=\frac{1}{\sigma_0^2}I_m\otimes\left(I_n-\frac{\tau_0^2}{n\tau_0^2+\sigma_0^2}J_n\right)-\frac{1}{mn(n\tau_0^2+\sigma_0^2)}J_m\otimes J_n.$$

From this, it can be shown that

$$\Pi_{11}=-\mathrm{tr}\{\mathcal{P}(\theta_0)(ZZ')\mathcal{P}(\theta_0)(ZZ')\}=-\frac{(m-1)n^2}{(n\tau_0^2+\sigma_0^2)^2}. \tag{S3.4}$$

As for Π_{12}, define

$$d_1=(1-p)^2+\frac{p^2}{m-1},\ d_2=\frac{2p(1-p)}{m-1}+\frac{(m-2)p^2}{(m-1)^2},$$

and note that $b'(\theta_0)b(\theta_0)=\tau_0^2\tilde{Z}\tilde{Z}'+\sigma_0^2I_N$, and

$$\tilde{Z}\tilde{Z}'=\begin{pmatrix}\tilde{Z}_1\tilde{Z}_1' & \cdots & \tilde{Z}_1\tilde{Z}_m' \\ & \cdots & \\ \tilde{Z}_m\tilde{Z}_1' & \cdots & \tilde{Z}_m\tilde{Z}_m'\end{pmatrix}.$$

Thus, we have $\mathrm{E}(\tilde{Z}_i\tilde{Z}_i')=(1-d_1)I_n+d_1J_n$ and $\mathrm{E}(\tilde{Z}_i\tilde{Z}_j')=d_2J_n, i\neq j$, implying that

$$\begin{aligned}\mathrm{E}\{b'(\theta_0)b(\theta_0)\} &= \tau_0^2\mathrm{E}(\tilde{Z}\tilde{Z}')+\sigma_0^2I_N\\ &= \tau_0^2\begin{pmatrix}d_1 & \cdots & d_2\\ & \cdots & \\ d_2 & \cdots & d_1\end{pmatrix}\otimes J_n+(1-d_1)\tau_0^2I_m\otimes I_n+\sigma_0^2I_m\otimes I_n\\ &\equiv \tau_0^2\mathcal{D}_m\otimes J_n+\{(1-d_1)\tau_0^2+\sigma_0^2\}I_m\otimes I_n. \end{aligned}\tag{S3.5}$$

Furthermore, we have the following expression:

$$\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)=\frac{n}{(n\tau_0^2+\sigma_0^2)^3}\Big(I_m-\frac{1}{m}J_m\Big)\otimes J_n. \tag{S3.6}$$

It is then easy to show that

$$\begin{aligned}
\Pi_{12} &= 2\mathrm{tr}\big[\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\mathrm{E}\{b'(\theta_0)b(\theta_0)\}\big] \\
&= \frac{2n}{(n\tau_0^2+\sigma_0^2)^3}\mathrm{tr}\bigg\{n\tau_0^2\mathcal{D}_m \otimes J_n - \frac{n\tau_0^2}{m}J_m \otimes J_n \\
&\quad +(1-d_1)\tau_0^2\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n + \sigma_0^2\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n\bigg\} \\
&= \frac{2n}{(n\tau_0^2+\sigma_0^2)^3}\{n^2\tau_0^2(md_1-1)+(m-1)(1-d_1)n\tau_0^2 \\
&\quad +(m-1)n\sigma_0^2\}. \qquad \text{(S3.7)}
\end{aligned}$$

Combining (S3.3), (S3.4), (S3.7), we have

$$\|V_1\|_{\mathrm{R}}^2 \sim \mathrm{E}_{\theta_0}\left[\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_1^2}\bigg|_{\theta=\theta_0}\right] = O(m).$$

Similarly, for $\|V_2\|_{\mathrm{R}}^2$, we have

$$\begin{aligned}
&\|V_2\|_{\mathrm{R}}^2 \sim \mathrm{E}_{\theta_0}\left[\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_2^2}\bigg|_{\theta=\theta_0}\right] \\
&= -\mathrm{tr}\big\{\mathcal{P}^2(\theta_0)\big\} + 2\mathrm{tr}\{\mathcal{P}^3(\theta_0)\mathrm{E}\{b'(\theta_0)b(\theta_0)\}\big] \\
&\equiv \Pi_{21} + \Pi_{22}. \qquad \text{(S3.8)}
\end{aligned}$$

For Π_{21}, using an expression above (S3.4), we have

$$\begin{aligned}
&\mathcal{P}^2(\theta_0) = \Big\{\mathcal{V}^{-1}(\theta_0) - \frac{1}{mn(n\tau_0^2+\sigma_0^2)}(J_m\otimes J_n)\Big\}^2 \\
&= \frac{1}{\sigma_0^4}I_m \otimes \Big\{I_n - \frac{n\tau_0^4+2\tau_0^2\sigma_0^2}{(n\tau_0^2+\sigma_0^2)^2}J_n\Big\} - \frac{1}{mn(n\tau_0^2+\sigma_0^2)^2}(J_m\otimes J_n).
\end{aligned}$$

Thus, we obtain the following expression:

$$\Pi_{21} = \frac{1}{(n\tau_0^2+\sigma_0^2)^2} - \frac{\{(n^2-n)\tau_0^4+2(n-1)\tau_0^2\sigma_0^2+\sigma_0^4\}mn}{\sigma_0^4(n\tau_0^2+\sigma_0^2)^2}. \qquad \text{(S3.9)}$$

As for Π_{22}, it can be shown that, for any $k \geq 1$, we have

$$\mathcal{P}^k(\theta_0) = \frac{1}{\sigma_0^{2k}} I_m \otimes \left(I_n - \frac{1}{n}J_n\right) + \frac{1}{n(n\tau_0^2+\sigma_0^2)^k}\left(I_m - \frac{1}{m}J_m\right) \otimes J_n. \tag{S3.10}$$

By (S3.5) and (S3.10), the following expression can be obtained

$$\begin{aligned}
&\Pi_{22} = 2\mathrm{tr}[\mathcal{P}^3(\theta_0)\mathrm{E}\{b'(\theta_0)b(\theta_0)\}] \\
&= \frac{2}{\sigma_0^6}\mathrm{tr}[\mathrm{E}\{b'(\theta_0)b(\theta_0)\}] - \frac{2(n^2\tau_0^6 + 3n\tau_0^4\sigma_0^2 + 3\tau_0^2\sigma_0^4)}{\sigma_0^6(n\tau_0^2+\sigma_0^2)^3}\mathrm{tr}[(I_m \otimes J_n)\mathrm{E}\{b'(\theta_0)b(\theta_0)\}] \\
&- \frac{2}{mn(n\tau_0^2+\sigma_0^2)^3}\mathrm{tr}[(J_m \otimes J_n)\mathrm{E}\{b'(\theta_0)b(\theta_0)\}] \\
&= mn\left\{\frac{2}{\sigma_0^4} + \frac{2\tau_0^2}{\sigma_0^6}(1-d_1)\right\}\{1+o(1)\}.
\end{aligned} \tag{S3.11}$$

Combining (S3.8), (S3.9) and (S3.11), it follows that

$$\|V_2\|_{\mathrm{R}}^2 \sim \mathrm{E}_{\theta_0}\left[\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_3^2}\Big|_{\theta=\theta_0}\right] = O(mn).$$

Thus, without loss of generality (w.l.o.g.) , we can let $p_1(N) = \sqrt{m}$ and $p_2(N) = \sqrt{mn}$.

Next, by Lemma 7.2 of Jiang (1996), we need an expression of $I_N(\theta_0)$ such that

$$\left[\frac{1}{p_i(N)p_j(N)}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i\partial\theta_j}\Big|_{\theta=\theta_0}\right]_{1\leq i,j\leq 2} = I_N(\theta_0) + o_{\mathrm{P}}(1) \tag{S3.12}$$

with $\liminf \lambda_{\min}\{I_N(\theta_0)\} > 0$ and $\limsup \lambda_{\max}\{I_N(\theta_0)\} < \infty$.

For the diagonal elements in (S3.12), we have

$$\frac{1}{p_1^2(N)}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_1^2}\bigg|_{\theta=\theta_0} = -\frac{1}{m}\mathrm{tr}\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\} + \frac{2}{m}W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0).$$

Then, by (S3.4)–(S3.7), it can be shown that

$$\mathrm{E}\left\{\frac{1}{p_1^2(N)}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_1^2}\bigg|_{\theta=\theta_0}\right\} = \frac{1}{\tau_0^4} + o(1), \tag{S3.13}$$

as $m, n \to \infty$ and $p \to 0$. Furthermore, we have

$$\begin{aligned}
&\mathrm{Var}\left\{\frac{1}{p_1^2(N)}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_1^2}\bigg|_{\theta=\theta_0}\right\} \\
= &\frac{4}{m^2}\mathrm{Var}\Big\{W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0)\Big\} \\
= &\frac{4}{m^2}\mathrm{E}\Big\{W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0)\Big\}^2 \\
&-\frac{4}{m^2}\mathrm{E}^2\Big\{W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0)\Big\}.
\end{aligned}$$

By the law of iterated expectations, we have

$$\begin{aligned}
&\mathrm{E}\Big\{W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0)\Big\}^2 \\
= &\mathrm{E}\Big(\mathrm{E}\Big[\Big\{W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0)\Big\}^2\Big|\tilde{Z}\Big]\Big).
\end{aligned}$$

Thus, by properties of the normal quadratic forms, we have

$$\begin{aligned}
&\mathrm{E}\Big[\Big\{W_N'(\theta_0)b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)W_N(\theta_0)\Big\}^2\Big|\tilde{Z}\Big] \\
= &\mathrm{tr}^2\Big\{b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)\Big\} \\
&+2\mathrm{tr}\Big\{b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)\Big\}^2.
\end{aligned}$$

This, combined with (S3.6), lead to the following expression:

$$\begin{aligned}
&\mathrm{Var}\left\{\frac{1}{p_1^2(N)}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_1^2}\Big|_{\theta=\theta_0}\right\} \\
= \;& \frac{4}{m^2}\mathrm{E}\Big[\mathrm{tr}^2\Big\{b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)\Big\} \\
& +2\mathrm{tr}\Big\{b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)\Big\}^2\Big] \\
& -\frac{4}{m^2}\mathrm{E}^2\Big[\mathrm{tr}\Big\{b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)\Big\}\Big] \\
= \;& \frac{4n^2}{m^2(n\tau_0^2+\sigma_0^2)^6}\mathrm{Var}\Big(\mathrm{tr}\Big[b(\theta_0)\Big\{\Big(I_m-\frac{1}{m}J_m\Big)\otimes J_n\Big\}b'(\theta_0)\Big]\Big) \\
& +\frac{8}{m^2}\mathrm{E}\left[\mathrm{tr}\Big\{b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)\Big\}^2\right]. \qquad \text{(S3.14)}
\end{aligned}$$

We now evaluate the terms on the right side of (S3.14). For the first term, we can obtain that

$$\begin{aligned}
\mathrm{tr}\Big[b(\theta_0)\Big\{\Big(I_m-\frac{1}{m}J_m\Big)\otimes J_n\Big\}b'(\theta_0)\Big] \;=\;& (m-1)n\sigma_0^2 \\
& +\tau_0^2\mathrm{tr}\Big[\tilde{Z}'\Big\{\Big(I_m-\frac{1}{m}J_m\Big)\otimes J_n\Big\}\tilde{Z}\Big].
\end{aligned}$$

We then have the following expression:

$$\begin{aligned}
&\mathrm{Var}\Big(\mathrm{tr}\Big[b(\theta_0)\Big\{\Big(I_m-\frac{1}{m}J_m\Big)\otimes J_n\Big\}b'(\theta_0)\Big]\Big) \\
= \;& \tau_0^4\mathrm{Var}\left[\mathrm{tr}\{\tilde{Z}'(I_m\otimes J_n)\tilde{Z}\}-\frac{1}{m}\mathrm{tr}\{\tilde{Z}'(J_m\otimes J_n)\tilde{Z}\}\right] \\
\le \;& 2\tau_0^4\mathrm{Var}\left[\mathrm{tr}\{\tilde{Z}'(I_m\otimes J_n)\tilde{Z}\}\right]+\frac{2\tau_0^4}{m^2}\mathrm{Var}\left[\mathrm{tr}\{\tilde{Z}'(J_m\otimes J_n)\tilde{Z}\}\right]. \qquad \text{(S3.15)}
\end{aligned}$$

For the first term on the right side of (S3.14), we apply Lemma 1 to get

$$\mathrm{Var}\left[\mathrm{tr}\{\tilde{Z}'(I_m\otimes J_n)\tilde{Z}\}\right]=mn^3pO(1).$$

As for the second term on the right side of (S3.14), note that

$$\begin{aligned}\mathrm{tr}\{\tilde{Z}'(J_m \otimes J_n)\tilde{Z}\} &= \sum_{i=1}^{m} \mathrm{tr}(\tilde{Z}_i' J_n \tilde{Z}_i) + \sum_{i \neq j} \mathrm{tr}(\tilde{Z}_i' J_n \tilde{Z}_j) \\ &= \mathrm{tr}\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\} + \sum_{i \neq j} \mathrm{tr}(\tilde{Z}_i' J_n \tilde{Z}_j).\end{aligned}$$

By the proof of Lemma 2, it can be seen that

$$\mathrm{Var}\big\{\mathrm{tr}(\tilde{Z}_i' J_n \tilde{Z}_j)\big\} = \mathrm{Var}\left(\sum_{l=1}^{m} S_{il} S_{jl}\right) = \frac{n^3 p O(1)}{m-1}.$$

It can be shown, using the Cauchy-Schwarz inequality, and by Lemma 1, that

$$\begin{aligned}&\frac{1}{m^2}\mathrm{Var}\Big(\mathrm{tr}\{\tilde{Z}'(J_m \otimes J_n)\tilde{Z}\}\Big) \le \frac{2}{m^2}\mathrm{Var}\left[\mathrm{tr}\{\tilde{Z}'(I_m \otimes J_n)\tilde{Z}\}\right] \\ &+\frac{2}{m^2}\mathrm{Var}\left[\sum_{i \neq j} \mathrm{tr}(\tilde{Z}_i' J_n \tilde{Z}_j)\right] = mn^3 p O(1).\end{aligned}$$

It can be seen now the first term on the right side of (S3.14) is $o(1)$.

By similar arguments, it can be shown that the second term on the right side of (S3.14) is $o(1)$. Thus, combined with (S3.13), we have

$$\frac{1}{p_1^2(N)} \frac{\partial^2 l_\theta(\tilde{y})}{\partial \theta_1^2}\bigg|_{\theta=\theta_0} = \frac{1}{\tau_0^4} + o_{\mathrm{P}}(1). \tag{S3.16}$$

By similar arguments, it can be shown that

$$\frac{1}{p_2^2(N)} \frac{\partial^2 l_\theta(\tilde{y})}{\partial \theta_2^2}\bigg|_{\theta=\theta_0} = \frac{1}{\sigma_0^4} + o_{\mathrm{P}}(1), \tag{S3.17}$$

$$\frac{1}{p_1(N)p_2(N)} \frac{\partial^2 l_\theta(\tilde{y})}{\partial \theta_1 \partial \theta_2}\bigg|_{\theta=\theta_0} = o_{\mathrm{P}}(1). \tag{S3.18}$$

It follows that (S3.12) holds with $I_N(\theta_0) = \mathrm{diag}(\tau_0^{-4}, \sigma_0^{-4})$.

Next, according to Lemma 7.2 of Jiang (1996), we need to show that there is a sequence of positive numbers, $q_i(N)$ such that $p_i(N)q_i(N) \to \infty$, and

$$\frac{1}{p_i(N)p_j(N)p_k(N)} \sup_{\theta \in \Theta_N} \left| \frac{\partial^3 l_\theta(\tilde{y})}{\partial\theta_i \partial\theta_j \partial\theta_k} \right| \xrightarrow{\text{P}} 0, \quad 1 \le i, j, k \le 2, \tag{S3.19}$$

where $\Theta_N = \{|\theta_i - \theta_{i0}| < q_i(N), i = 1, 2\}$. Let $q_i(N) = \sqrt{p_1(N) \wedge p_2(N)}/p_i(N)$, $i = 1, 2$. It follows that $q_1(N) = m^{-1/4}$ and $q_2(N) = m^{-1/4}n^{-1/2}$.

For $i = j = k = 1$ in (S3.19), we begin with the following expression:

$$\begin{aligned} \frac{\partial^3 l_\theta(\tilde{y})}{\partial\theta_1^3} &= 2\text{tr}\big\{\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ')\big\} \\ &\quad -6W_N'(\theta_0)b(\theta_0)\big\{\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)\big\}b'(\theta_0)W_N(\theta_0). \end{aligned}$$

It follows that $\sup_{\theta \in \Theta_N} |\partial^3 l_\theta(\tilde{y})/\partial\theta_1^3| \le 2I_1 + 6I_2$, where

$$\begin{aligned} I_1 &= \sup_{\theta \in \Theta_N} \text{tr}\{\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\}, \\ I_2 &= \sup_{\theta \in \Theta_N} |W_N'(\theta_0)b(\theta_0)\{\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)\}b'(\theta_0)W_N(\theta_0)|. \end{aligned}$$

For I_1, note that we have the following expression:

$$\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ') = \frac{n^2}{(n\tau^2 + \sigma^2)^3}\Big(I_m - \frac{1}{m}J_m\Big) \otimes J_n.$$

It can then be shown that $p_1^{-3}(N)\text{E}(I_1) = O(1)/\sqrt{m}$. Similarly, for I_2, we have

$$\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta)(ZZ')\mathcal{P}(\theta) = \frac{n^2}{(n\tau^2 + \sigma^2)^4}\Big(I_m - \frac{1}{m}J_m\Big) \otimes J_n.$$

It follows that, for any $\theta \in \Theta_N$, we have

$$\begin{aligned}
& W_N'(\theta_0)b(\theta_0)\{\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)ZZ'\mathcal{P}(\theta)\}b'(\theta_0)W_N(\theta_0) \\
\leq\ & \frac{c}{n^2}W_N'(\theta_0)b(\theta_0)\Big\{\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n\Big\}b'(\theta_0)W_N(\theta_0) \\
=\ & \frac{c}{n^2}\mathrm{tr}\Big[\Big\{\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n\Big\}(\epsilon\epsilon' + \tilde{Z}\alpha\alpha'\tilde{Z}')\Big],
\end{aligned}$$

for some constant c, noting that $b'(\theta_0)W_N(\theta_0)W_N'(\theta_0)b(\theta_0) = \epsilon\epsilon' + \tilde{Z}\alpha\alpha'\tilde{Z}'$. Also note that $\mathrm{E}(\tilde{Z}\alpha\alpha'\tilde{Z}') = \mathrm{E}\{\mathrm{E}(\tilde{Z}\alpha\alpha'\tilde{Z}'|\tilde{Z})\} = \tau_0^2\mathrm{E}(\tilde{Z}\tilde{Z}')$, and, according to an earlier result [see (S3.5)], $\mathrm{E}(\tilde{Z}\tilde{Z}') = \mathcal{D}_m \otimes J_n + (1-d_1)(I_m \otimes I_n)$. It follows

$$\mathrm{E}(I_2) \leq \frac{c}{n^2}\mathrm{E}\left(\mathrm{tr}\Big[\Big\{\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n\Big\}\{\sigma_0^2 I_N + \tau_0^2\mathrm{E}(\tilde{Z}\tilde{Z}')\}\Big]\right) = O(m).$$

Thus, we have $p_1^{-3}(N)\mathrm{E}(I_2) = O(1)/\sqrt{m}$, which, combined with the earlier results, imply that (S3.19) holds for $i = j = k = 1$. By similar arguments, it can be shown that (S3.19) holds for other combinations of i, j, k.

Finally, let $A_N(\theta_0) = [A_{N1}(\theta_0), A_{N2}(\theta_0)]' =$

$$\left[\frac{1}{p_1(N)}\left.\frac{\partial l_\theta(\tilde{y})}{\partial\theta_1}\right|_{\theta=\theta_0}, \frac{1}{p_2(N)}\left.\frac{\partial l_\theta(\tilde{y})}{\partial\theta_2}\right|_{\theta=\theta_0}\right]'.$$

By Lemma 7.2 in Jiang (1996), we need to show boundedness of $\{|A_N(\theta_0)|\}$. First, write

$$\begin{aligned}
& A_{N1}(\theta_0) = \frac{1}{\sqrt{m}}\left.\frac{\partial l_\theta(\tilde{y})}{\partial\tau^2}\right|_{\theta=\theta_0} \\
& = \frac{1}{\sqrt{m}}\mathrm{tr}\{\mathcal{P}(\theta_0)ZZ'\} - \frac{1}{\sqrt{m}}W_N'(\theta_0)b(\theta_0)\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\}b'(\theta_0)W_N(\theta_0) \\
& \equiv \Pi_{31} - \Pi_{32} = \{\Pi_{31} - \mathrm{E}(\Pi_{32})\} - \{\Pi_{32} - \mathrm{E}(\Pi_{32})\}. \qquad \text{(S3.20)}
\end{aligned}$$

For Π_{31}, we have $\mathcal{P}(\theta_0)ZZ' = (n\tau_0^2 + \sigma_0^2)^{-1}(I_m - m^{-1}J_m) \otimes J_n$, implying

$$\Pi_{31} = \frac{n(m-1)}{\sqrt{m}(n\tau_0^2 + \sigma_0^2)}.$$

Next, we have $\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0) = (n\tau_0^2 + \sigma_0^2)^{-1}(I_m - m^{-1}J_m) \otimes J_n$, implying

$$\mathrm{E}(\Pi_{32}) = \frac{n^2(md_1 - 1)\tau_0^2 + n(m-1)\{(1-d_1)\tau_0^2 + \sigma_0^2\}}{\sqrt{m}(n\tau_0^2 + \sigma_0^2)^2}.$$

By the definition of d_1 [see below (S3.4)], it can then be shown that

$$\Pi_{31} - \mathrm{E}(\Pi_{32}) = \sqrt{m}\left(p + \frac{1}{n}\right)O(1) = \frac{\sqrt{m}}{n}O(1). \qquad \text{(S3.21)}$$

On the other hand, by following the same arguments as those leading to (S3.16), and applying Lemma 1 and Lemma 2, it can be shown that

$$\mathrm{Var}(\Pi_{32}) = \mathrm{E}\{\Pi_{32} - \mathrm{E}(\Pi_{32})\}^2 = O(1). \qquad \text{(S3.22)}$$

By (S3.21), (S3.22), and the fact that $m \sim n$, implied by the condition of the theorem, it follows that $|A_{N1}(\theta_0)| = O_{\mathrm{P}}(1)$. Similarly, define $\Pi_{41} = \mathrm{tr}\{\mathcal{P}(\theta_0)\}/\sqrt{mn}$, and $\Pi_{42} = W_N'(\theta_0)b(\theta_0)\mathcal{P}^2(\theta_0)b'(\theta_0)W_N(\theta_0)/\sqrt{mn}$. By similar arguments, it can be shown that

$$\Pi_{41} - \mathrm{E}(\Pi_{42}) = \sqrt{mn}p + o(1), \quad \Pi_{42} - \mathrm{E}(\Pi_{42}) = O_{\mathrm{P}}(1). \qquad \text{(S3.23)}$$

Thus, by similar arguments, it can be shown that $|A_{N2}(\theta_0)| = O_{\mathrm{P}}(1)$.

The result then follows by applying Lemma 7.2 (i) of Jiang (1996).

S4 Proof of Theorem 2

We provide the proof for the REML part. Following the notation and proof of Theorem 1, in particular, with $p_1(N) = \sqrt{m}$ and $p_2(N) = \sqrt{mn}$, define $\Sigma(\theta_0) = 2\text{diag}(\tau_0^{-4}, \sigma_0^{-4})$. For any $a \in \mathrm{R}^2 \setminus \{0\}$, we have

$$a'\Sigma^{-1/2}(\theta_0)(A_{N1}(\theta_0), A_{N2}(\theta_0))' = a'\Sigma^{-1/2}(\theta_0)(A_{N11}(\theta_0), A_{N21}(\theta_0))'$$
$$-a'\Sigma^{-1/2}(\theta_0)(A_{N12}(\theta_0), A_{N22}(\theta_0))', \quad \text{(S4.1)}$$

where $A_{N11}(\theta_0) = \Pi_{31} - \mathrm{E}(\Pi_{32})$, $A_{N12}(\theta_0) = \Pi_{32} - \mathrm{E}(\Pi_{32})$, $A_{N21}(\theta_0) = \Pi_{41} - \mathrm{E}(\Pi_{42})$, $A_{N22}(\theta_0) = \Pi_{42} - \mathrm{E}(\Pi_{42})$, $A_{N11}(\theta_0)$ and $A_{N21}(\theta_0)$ represent the bias of the REML estimator, while $A_{N12}(\theta_0)$ and $A_{N22}(\theta_0)$ the volatility of the estimator under the mis-classified LMM. By (S3.21) and (S3.23), and the condition of the theorem, it is seen that the bias terms asymptotically vanish as $m, n \to \infty$. Thus, we can focus on the volatility terms.

Let $c_a = (c_{a1}, c_{a2}) = \Sigma^{-1/2}(\theta_0)a$, $\Delta_N(\theta_0) = (c_{a1}/\sqrt{m})V_1(\theta_0) + (c_{a2}/\sqrt{mn})V_2(\theta_0)$, where $V_1(\theta_0) = b(\theta_0)\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)b'(\theta_0)$ and $V_2(\theta_0) = b(\theta_0)\mathcal{P}^2(\theta_0)b'(\theta_0)$. We have

$$a'\Sigma^{-1/2}(\theta_0)[A_{N12}(\theta_0), A_{N22}(\theta_0)]' = c_a'[A_{N12}(\theta_0), A_{N22}(\theta_0)]'$$
$$= W_N'(\theta_0)\big[\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\}\big]W_N(\theta_0)$$
$$+W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0) - \mathrm{E}\{W_N'(\theta_0)\Delta_N(\theta_0)W_N(\theta_0)\}$$
$$\equiv \Pi_{51} + \Pi_{52}, \quad \text{(S4.2)}$$

with Π_{51}, Π_{52} defined in obvious ways.

We first show that $\Pi_{51} = o_{\mathrm{P}}(1)$. By properties of quadratic forms of random variables, we have $\mathrm{E}(\Pi_{51}|\tilde{Z}) = \mathrm{tr}[\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\}]$. It follows that

$$\begin{aligned}\mathrm{E}\{\mathrm{E}^2(\Pi_{51}|\tilde{Z})\} &= \mathrm{E}\big(\mathrm{tr}\{\Delta_N(\theta_0)\} - \mathrm{E}[\mathrm{tr}\{\Delta_N(\theta_0)\}]\big)^2 \\ &\leq \frac{c_{a1}^2}{m}\mathrm{Var}\big[\mathrm{tr}\{V_1(\theta_0)\}\big] + \frac{c_{a2}^2}{mn}\mathrm{Var}\big[\mathrm{tr}\{V_2(\theta_0)\}\big] \\ &\quad +2c_{a1}c_{a2}\sqrt{\frac{\mathrm{Var}\big[\mathrm{tr}\{V_1(\theta_0)\}\big]}{m}}\sqrt{\frac{\mathrm{Var}\big[\mathrm{tr}\{V_2(\theta_0)\}\big]}{mn}}.\end{aligned}$$

It can be shown that $\mathrm{tr}\{V_1(\theta_0)\} =$

$$\frac{\sigma_0^2}{(n\tau_0^2+\sigma_0^2)^2}\mathrm{tr}\Big\{\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n\Big\} + \frac{\tau_0^2}{(n\tau_0^2+\sigma_0^2)^2}\mathrm{tr}\Big[\tilde{Z}'\Big\{\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n\Big\}\tilde{Z}\Big].$$

Thus, by applying Lemma 1, it can be shown that $\mathrm{Var}\big[\mathrm{tr}\{V_1(\theta_0)\}\big]/m = o(1)$. Similarly, it can be shown that $\mathrm{Var}\big[\mathrm{tr}\{V_2(\theta_0)\}\big]/mn = o(1)$. It follows that $\mathrm{E}\{\mathrm{E}^2(\Pi_{51}|\tilde{Z})\} = o(1)$, implying $\mathrm{E}(\Pi_{51}|\tilde{Z}) = o_{\mathrm{P}}(1)$.

Next, by properties of quadratic forms of normal random variables, we have

$$\mathrm{E}\{\mathrm{Var}(\Pi_{51}|\tilde{Z})\} = 2\mathrm{E}(\mathrm{tr}[\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\}]^2).$$

Furthermore, write $\Omega(\theta_0) = \mathcal{P}(\theta_0)(ZZ')\mathcal{P}(\theta_0)$ and

$$\tilde{\Delta}_N(\theta_0) = (c_{a1}/\sqrt{m})\Omega(\theta_0) + (c_{a2}/\sqrt{mn})\mathcal{P}^2(\theta_0).$$

The following expression can then be derived: $\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\} =$

$$\begin{bmatrix} 0 & \sigma_0\tau_0\tilde{\Delta}_N(\theta_0)\{\tilde{Z} - \mathrm{E}(\tilde{Z})\} \\ \sigma_0\tau_0\{\tilde{Z}' - \mathrm{E}(\tilde{Z}')\}\tilde{\Delta}_N(\theta_0) & \tau_0^2\big[\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z} - \mathrm{E}\{\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z}\}\big] \end{bmatrix}.$$

It then follows that $\mathrm{E}\{\mathrm{Var}(\Pi_{51}|\tilde{Z})\} = 2\mathrm{E}(\mathrm{tr}[\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\}]^2) =$

$$4\sigma_0^2\tau_0^2\mathrm{tr}\big[\tilde{\Delta}_N(\theta_0)\{\mathrm{E}(\tilde{Z}\tilde{Z}') - \mathrm{E}(\tilde{Z})\mathrm{E}(\tilde{Z}')\}\tilde{\Delta}_N(\theta_0)\big]$$
$$+2\tau_0^4\mathrm{E}\Big(\mathrm{tr}\big[\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z} - \mathrm{E}\{\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z}\}\big]^2\Big).$$

By (S3.5), and the fact that $\mathrm{E}(\tilde{Z})\mathrm{E}(\tilde{Z}') = \mathcal{D}_m \otimes J_n$, we have

$$\mathrm{tr}\big[\tilde{\Delta}_N(\theta_0)\{\mathrm{E}(\tilde{Z}\tilde{Z}') - \mathrm{E}(\tilde{Z})\mathrm{E}(\tilde{Z}')\}\tilde{\Delta}_N(\theta_0)\big] = (1-d_1)\mathrm{tr}\big\{\tilde{\Delta}_N^2(\theta_0)\big\} = o(1).$$

Furthermore, the following expressions can be derived:

$$\mathrm{E}(\mathrm{tr}\big[\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z} - \mathrm{E}\{\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z}\}]^2)$$
$$= \mathrm{E}[\mathrm{tr}\{\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z}\}^2] - \mathrm{tr}([\mathrm{E}\{\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z}\}]^2),$$
$$\tilde{Z}'\Omega(\theta_0)\tilde{Z} = \frac{1}{(n\tau_0^2+\sigma_0^2)^2}\tilde{Z}'\Big\{\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n\Big\}\tilde{Z},$$
$$\tilde{Z}'\mathcal{P}^2(\theta_0)\tilde{Z} = \tilde{Z}'\left\{\frac{1}{\sigma_0^4}I_m\otimes I_n - \frac{n\tau_0^4+2\tau_0^2\sigma_0^2}{\sigma_0^4(n\tau_0^2+\sigma_0^2)^2}I_m\otimes J_n\right.$$
$$\left. - \frac{1}{mn(n\tau_0^2+\sigma_0^2)^2}J_m\otimes J_n\right\}\tilde{Z}.$$

Thus, by applying the arguments as used in the proof of Lemma 2, it can be shown that

$$\mathrm{E}\Big(\mathrm{tr}\big[\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z} - \mathrm{E}\{\tilde{Z}'\tilde{\Delta}_N(\theta_0)\tilde{Z}\}\big]^2\Big) = o(1).$$

Combining the above results, we have $\mathrm{E}\{\mathrm{Var}(\Pi_{51}|\tilde{Z})\} = o(1)$.

The established results imply that $\mathrm{E}(\Pi_{51}^2|\tilde{Z}) = o_{\mathrm{P}}(1)$. It follows by the dominated convergence theorem that $\Pi_{51} = o_{\mathrm{P}}(1)$. Thus, we can further focus on the term Π_{52}.

By the assumption of the theorem, we have

$$\Pi_{52} = W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0) - \mathrm{E}[W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0)].$$

Note that $W_N(\theta_0) \sim N(0, I_{N+m})$. Thus, by properties of quadratic forms of normal random variables, we have $\mathrm{Var}(\Pi_{52}) = 2\mathrm{tr}[\mathrm{E}^2\{\Delta_N(\theta_0)\}] =$

$$\frac{2}{m}\left(c_{a1}^2\mathrm{tr}\Big[\mathrm{E}^2\{V_1(\theta_0)\}\Big] + 2\frac{c_{a1}c_{a2}}{\sqrt{n}}\mathrm{tr}\Big[\mathrm{E}\{V_1(\theta_0)\}\mathrm{E}\{V_2(\theta_0)\}\Big] + \frac{c_{a2}^2}{n}\mathrm{tr}\Big[\mathrm{E}^2\{V_2(\theta_0)\}\Big]\right).$$

It can be shown that $\mathrm{tr}[\mathrm{E}\{V_1(\theta_0)\}\mathrm{E}\{V_1(\theta_0)\}] = U_1/(n\tau_0^2+\sigma_0^2)^4$, where

$$U_1 = \sigma_0^4(m-1)n^2 + \tau_0^4 n^4[m\{d_1^2 + (m-1)d_2^2\} - 1] + 2\sigma_0^2\tau_0^2 n^3(md_1 - 1).$$

It then follows that $m^{-1}\mathrm{tr}[\mathrm{E}^2\{V_1(\theta_0)\}] \to \tau_0^{-4}$. Next, we have

$$\begin{aligned}
&\mathrm{tr}\Big[\mathrm{E}\{V_1(\theta_0)\}\mathrm{E}\{V_2(\theta_0)\}\Big] = \sigma_0^4\mathrm{tr}\big[\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\}\mathcal{P}^2(\theta_0)\big]\\
&+\tau_0^4\mathrm{tr}\big[\mathrm{E}\{\tilde{Z}'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\tilde{Z}\}\mathrm{E}\{\tilde{Z}'\mathcal{P}^2(\theta_0)\tilde{Z}\}\big]\\
&+2\sigma_0^2\tau_0^2\mathrm{tr}\big[\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\}\mathrm{E}(\tilde{Z})\mathrm{E}(\tilde{Z}')\mathcal{P}^2(\theta_0)\big].
\end{aligned}$$

By similar arguments to those used above, the following expressions can be derived:

$$\begin{aligned}
&\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\}\mathcal{P}^2(\theta_0) = \frac{1}{(n\tau_0^2+\sigma_0^2)^4}\Big(I_m - \frac{1}{m}J_m\Big)\otimes J_n,\\
&\mathrm{E}\{\tilde{Z}'\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\tilde{Z}\}\mathrm{E}\{\tilde{Z}'\mathcal{P}^2(\theta_0)\tilde{Z}\}\\
&= \frac{n^3}{\sigma_0^4(n\tau_0^2+\sigma_0^2)^2}\left[\mathcal{D}_m - \frac{1}{m}J_m - \frac{n^2\tau_0^4 + 2n\tau_0^2\sigma_0^2}{(n\tau_0^2+\sigma_0^2)^2}\{d_1^2 + (m-1)d_2^2 - m^{-1}\}J_m\right],\\
&\{\mathcal{P}(\theta_0)ZZ'\mathcal{P}(\theta_0)\}\mathrm{E}(\tilde{Z})\mathrm{E}(\tilde{Z}')\mathcal{P}^2(\theta_0) = \frac{n}{(n\tau_0^2+\sigma_0^2)^4}\Big(\mathcal{D}_m - \frac{1}{m}J_m\Big)\otimes J_n.
\end{aligned}$$

It follows, by the condition of the theorem, that $(m\sqrt{n})^{-1}\mathrm{tr}[\mathrm{E}\{V_1(\theta_0)\}\mathrm{E}\{V_2(\theta_0)\}] =$

$$\frac{1}{m\sqrt{n}}\left\{\frac{\sigma_0^4(m-1)n}{(n\tau_0^2+\sigma_0^2)^4}+\frac{2\sigma_0^2\tau_0^2n^2(md_1-1)}{(n\tau_0^2+\sigma_0^2)^4}+\frac{n^3\tau_0^4}{\sigma_0^4(n\tau_0^2+\sigma_0^2)^2}\left[md_1-1\right.\right.$$
$$\left.\left.-\frac{n^2\tau_0^4+2n\tau_0^2\sigma_0^2}{(n\tau_0^2+\sigma_0^2)^2}\{md_1^2+m(m-1)d_2^2-1\}\right]\right\}=O(\sqrt{n}p)=o(1).$$

Finally, we have $\mathrm{tr}[\mathrm{E}^2\{V_2(\theta_0)\}] =$

$$\mathrm{tr}\left[\sigma_0^4\mathcal{P}^4(\theta_0)+2\sigma_0^2\tau_0^2\mathcal{P}^2(\theta_0)\mathrm{E}(\tilde{Z})\mathrm{E}(\tilde{Z}')\mathcal{P}^2(\theta_0)+\tau_0^4\mathrm{E}^2\{\tilde{Z}'\mathcal{P}^2(\theta_0)\tilde{Z}\}\right].$$

It can be derived that $\mathcal{P}^4(\theta_0)=\sigma_0^{-8}I_m\otimes[I_n-\lambda_1J_n]+\lambda_2J_m\otimes J_n$ with

$$\lambda_1=\frac{2(n\tau_0^4+2\tau_0^2\sigma_0^2)}{(n\tau_0^2+\sigma_0^2)^2}-\frac{n(n\tau_0^4+2\tau_0^2\sigma_0^2)^2}{(n\tau_0^2+\sigma_0^2)^4},\quad \lambda_2=\frac{(n\tau_0^2+\sigma_0^2-2)}{mn(n\tau_0^2+\sigma_0^2)^3}.$$

It can then be shown that $(\sigma_0^4/mn)\mathrm{tr}\{\mathcal{P}^4(\theta_0)\}\to\sigma_0^{-4}$. By similar arguments, it can be shown that $\mathrm{tr}\{\mathcal{P}^2(\theta_0)\mathrm{E}(\tilde{Z})\mathrm{E}(\tilde{Z}')\mathcal{P}^2(\theta_0)\}/mn$ and $\mathrm{tr}[\mathrm{E}^2\{\tilde{Z}'\mathcal{P}^2(\theta_0)\tilde{Z}\}]/mn$ are both $o(1)$. It follows that $\mathrm{tr}[\mathrm{E}^2\{V_2(\theta_0)\}]/mn\to\sigma_0^{-4}$. Combining the above results, we have

$$\mathrm{Var}(\Pi_{52})\to c_a'\Sigma(\theta_0)c_a=|a|^2,$$

hence $\mathrm{tr}[\mathrm{E}^2\{\Delta_N(\theta_0)\}]=\mathrm{Var}(\Pi_{52})/2\to|a|^2/2>0$.

On the other hand, it can be shown that there are constants $c_1, c_2>0$ such that

$$\lambda_{\max}\left[\mathrm{E}^2\{\Delta_N(\theta_0)\}\right]\le\left(\frac{c_1}{m}+\frac{c_2}{mn}\right)\frac{|a|^2}{\lambda_{\min}\{\Sigma(\theta_0)\}}.$$

Combing the above results, we have

$$\frac{\lambda_{\max}\left[\mathrm{E}^2\{\Delta_N(\theta_0)\}\right]}{\mathrm{tr}\left[\mathrm{E}^2\{\Delta_N(\theta_0)\}\right]}\le\frac{2}{\lambda_{\min}\{\Sigma(\theta_0)\}}\left(\frac{c_1}{m}+\frac{c_2}{mn}\right)\longrightarrow 0. \tag{S4.3}$$

Therefore, by applying Theorem 5.1 of Jiang (1996), noting that a key condition, (S4.3), has been verified, we conclude that

$$\frac{W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0) - \mathrm{E}\{W_N'(\theta_0)\Delta_N(\theta_0)W_N(\theta_0)\}}{\sqrt{\mathrm{Var}\Big[W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0)\Big]}} \xrightarrow{\mathrm{d}} N(0,1).$$

The asymptotic normality of the REML estimator follows by the arbitrariness of a.

S5 Proof of Theorem 3

By Lemma 7.2 in Jiang (1996), we need to first identify the sequences $p_l(N)$, $1 \le l \le s+1$. By Lemma 4.2 in Jiang (1996) and the AI^4 condition, we have $p_l(N) \sim \|V_l\|_{\mathrm{R}}$ and

$$\|V_l\|_{\mathrm{R}}^2 \sim \mathrm{E}_{\theta_0}\left[\left.\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_l^2}\right|_{\theta=\theta_0}\right], \quad 1 \le l \le s+1.$$

Note that, here, $\theta \equiv (\theta_1, \cdots, \theta_{s+1})' = (\tau_1^2, \cdots, \tau_s^2, \sigma^2)'$ and $\theta_0 = (\theta_{10}, \cdots, \theta_{s+1,0})'$ $= (\tau_{10}^2, \cdots, \tau_{s0}^2, \sigma_0^2)'$. Thus, it suffices to obtain the asymptotic orders of $\|V_l\|_{\mathrm{R}}$, $1 \le l \le s+1$.

By matrix differentiation, it is easy to obtain the following expressions:

$$\begin{aligned}\frac{\partial l_\theta(\tilde{y})}{\partial\theta_l} &= \frac{\partial \ln|V(\Phi,\theta)|}{\partial\theta_l} + \frac{\partial\{\tilde{y}'V^{-1}(\Phi,\theta)\tilde{y}\}}{\partial\theta_l} \\ &= \mathrm{tr}\left\{V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}\right\} - \tilde{y}'V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\tilde{y}\end{aligned}$$

with $\partial V(\Phi,\theta)/\partial\theta_l = \Phi' U_l U_l' \Phi$ for $1 \le l \le s$ and $\partial V(\Phi,\theta)/\partial\theta_{s+1} = \Phi'\Phi$. Thus, we have

$$\begin{aligned}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_l^2} &= -\mathrm{tr}\Big\{V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}\Big\} \\ &\quad +2\tilde{y}'V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\tilde{y}.\end{aligned}$$

For the mixed ANOVA model (3.5) in MJ22, it follows that $\tilde{y} = \Phi'\tilde{U}_1\alpha_1 + \cdots + \Phi'\tilde{U}_s\alpha_s + \Phi'\epsilon$. Define $b(\theta) = (\sigma I_N, \tau_1\tilde{U}_1, \cdots, \tau_s\tilde{U}_s)'$, and $W_N(\theta) = \big(W_{N1}(\theta), \cdots, W_{N,s+1}(\theta)\big)'$ with $W_{Nl}(\theta) = \alpha_l/\tau_l$ for $1 \le l \le s$ and $W_{Nl}(\theta) = \epsilon/\sigma$ for $l = s+1$. By the assumptions of model (3.5) and assumption (i) in Theorem 5, it follows that $W_{Ni}(\theta_0)$'s are independent with $b(\theta_0)$, and are distributed as standard multivariate normal, $b'(\theta_0)b(\theta_0) = \sum_{k=1}^s \tau_{k0}^2\tilde{U}_k'\tilde{U}_k + I_N\sigma_0^2$, $\tilde{y} = \Phi' b'(\theta_0)W_N(\theta_0)$. Then, we have

$$\begin{aligned}&\mathrm{E}_{\theta_0}\Big[\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_l^2}\Big|_{\theta=\theta_0}\Big] \\ =\ & -\mathrm{tr}\Big\{V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}\Big\} \\ & +2\mathrm{E}_{\theta_0}\Big[\tilde{y}'V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}V^{-1}(\Phi,\theta_0)\tilde{y}\Big] \\ =\ & -\mathrm{tr}\Big\{V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}\Big\} + 2\mathrm{E}_{\theta_0}\Big[W_N'(\theta_0)b(\theta_0)\Phi \\ & V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_l}V^{-1}(\Phi,\theta_0)\Phi' b'(\theta_0)W_N(\theta_0)\Big] \\ \equiv\ & -\mathrm{tr}\{\Psi_{2ll}(\theta_0)\} + 2\mathrm{E}_{\theta_0}\big\{W_N'(\theta_0)A_{2ll}(\theta_0)W_N(\theta_0)\big\} \\ =\ & 2\mathrm{E}_{\theta_0}[\mathrm{tr}\{A_{2ll}(\theta_0)\}] - \mathrm{tr}\{\Psi_{2ll}(\theta_0)\},\end{aligned}$$

with $A_{2ll}(\theta_0) = b(\theta_0)\Phi\Psi_{2ll}(\theta_0)V^{-1}(\Phi,\theta_0)\Phi' b'(\theta_0)$. Thus, we have

$$p_l(N) \sim \|V_l\|_{\mathrm{R}} = \sqrt{2\mathrm{E}_{\theta_0}[\mathrm{tr}\{A_{2ll}(\theta_0)\}] - \mathrm{tr}\{\Psi_{2ll}(\theta_0)\}}, l = 1, \cdots, s+1.$$

Next, by Lemma 7.2 in Jiang (1996), we need to obtain a detailed expression of $I_N(\theta_0)$ with $\liminf \lambda_{\min}\{I_N(\theta_0)\} > 0$ such that

$$\left[\frac{1}{p_i(N)p_j(N)} \frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i \partial\theta_j}\Big|_{\theta=\theta_0}\right]_{1\le i,j\le s+1} = I_N(\theta_0) + o_p(1). \tag{S5.1}$$

For the diagonal elements in (S5.1), by assumption (ii), we have

$$\begin{aligned}
& \mathrm{E}_{\theta_0}\left\{\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i^2}\Big|_{\theta=\theta_0}\right\} \\
= \;& -\mathrm{tr}\{\Psi_{2ii}(\theta_0)\} + 2\mathrm{E}_{\theta_0}[\mathrm{tr}\{A_{2ii}(\theta_0)\}] \\
= \;& \mathrm{tr}\{\Psi_{2ii}(\theta_0)\} + 2\mathrm{E}_{\theta_0}\big[\mathrm{tr}\{A_{2ii}(\theta_0) - \Psi_{2ii}(\theta_0)\}\big] \\
= \;& \mathrm{tr}\{\Psi_{2ii}(\theta_0)\} + 2\mathrm{E}_{\theta_0}\left\{\mathrm{tr}\big(\Psi_{2ii}(\theta_0)V^{-1}(\Phi,\theta_0)\big[\Phi'\mathrm{E}_{\theta_0}\{b'(\theta_0)b(\theta_0)\}\Phi - V(\Phi,\theta_0)\big]\big)\right\} \\
= \;& \{\Psi_{2ii}(\theta_0)\}\{1 + o(1)\}.
\end{aligned}$$

Also, by assumptions (i) and (ii) of Theorem 3, and properties of the multivariate normal distribution (e.g., Jiang and Nguyen (2021), sec. B.1), we have

$$\begin{aligned}
& \mathrm{Var}_{\theta_0}\left\{\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i^2}\Big|_{\theta=\theta_0}\right\} \\
= \;& \mathrm{Var}_{\theta_0}\{2W_N'(\theta_0)A_{2ii}(\theta_0)W_N(\theta_0)\}
\end{aligned}$$

$$
\begin{aligned}
&= 4\Big[\mathrm{E}_{\theta_0}\{W_N'(\theta_0)A_{2ii}(\theta_0)W_N(\theta_0)\}^2 - \mathrm{E}_{\theta_0}^2\{W_N'(\theta_0)A_{2ii}(\theta_0)W_N(\theta_0)\}\Big] \\
&= 4\Big\{\mathrm{E}_{\theta_0}\Big(\big[\mathrm{tr}\{A_{2ii}(\theta_0)\}\big]^2 + 2\mathrm{tr}\{A_{2ii}^2(\theta_0)\}\Big) - \mathrm{E}_{\theta_0}^2[\mathrm{tr}\{A_{2ii}(\theta_0)\}]\Big\} \\
&= 4\mathrm{Var}_{\theta_0}[\mathrm{tr}\{A_{2ii}(\theta_0)\}] + 8\mathrm{E}_{\theta_0}\Big(\mathrm{tr}\{A_{2ii}^2(\theta_0)\}\Big) \\
&= p_i^4(N)o(1).
\end{aligned}
$$

Thus, we have, for $l = 1, \ldots, s+1$,

$$
\frac{1}{p_i^2(N)}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i^2}\Big|_{\theta=\theta_0} = \frac{\mathrm{tr}\{\Psi_{2ii}(\theta_0)\}}{p_i^2(N)} + o_p(1).
$$

For the off-diagonal elements in (S5.1), by assumptions (i) and (iii), we have

$$
\begin{aligned}
&\mathrm{E}_{\theta_0}\left\{\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i\partial\theta_j}\Big|_{\theta=\theta_0}\right\} \\
&= -\mathrm{tr}\Big\{V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_i}V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_j}\Big\} \\
&\quad +2\mathrm{E}_{\theta_0}\Big\{\tilde{y}'V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_i}V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_j}V^{-1}(\Phi,\theta_0)\tilde{y}\Big\} \\
&\equiv -\mathrm{tr}\{\Psi_{2ij}(\theta_0)\} + 2\mathrm{E}_{\theta_0}\{W_N'(\theta_0)A_{2ij}(\theta_0)W_N(\theta_0)\} \\
&= -\mathrm{tr}\{\Psi_{2ij}(\theta_0)\} + 2\mathrm{E}_{\theta_0}[\mathrm{tr}\{A_{2ij}(\theta_0)\}] \\
&= \mathrm{tr}\{\Psi_{2ij}(\theta_0)\} + 2\mathrm{E}_{\theta_0}[\mathrm{tr}\{A_{2ij}(\theta_0) - \Psi_{2ij}(\theta_0)\}] \\
&= \mathrm{tr}\{\Psi_{2ij}(\theta_0)\}\{1 + o(1)\},
\end{aligned}
$$

where $A_{2ij}(\theta_0) = b(\theta_0)\Phi\Psi_{2ij}(\theta_0)V^{-1}(\Phi,\theta_0)\Phi'b'(\theta_0)$.

Also by assumptions (i) and (iii) and, again, the properties of multivariate normal distribution, it follows that

$$\begin{aligned}
&\mathrm{Var}_{\theta_0}\left\{\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i\partial\theta_j}\Big|_{\theta=\theta_0}\right\} \\
=\ & 4\Big[\mathrm{E}_{\theta_0}\{W_N'(\theta_0)A_{2ij}(\theta_0)W_N(\theta_0)\}^2 - \mathrm{E}_{\theta_0}^2\{W_N'(\theta_0)A_{2ij}(\theta_0)W_N(\theta_0)\}\Big] \\
=\ & 4\Big\{\mathrm{E}_{\theta_0}\big(\big[\mathrm{tr}\{A_{2ij}(\theta_0)\}\big]^2 + 2\mathrm{tr}\{A_{2ij}^2(\theta_0)\}\big) - \mathrm{E}_{\theta_0}^2[\mathrm{tr}\{A_{2ij}(\theta_0)\}]\Big\} \\
=\ & 4\mathrm{Var}_{\theta_0}\big[\mathrm{tr}\{A_{2ij}(\theta_0)\}\big] + 8\mathrm{E}_{\theta_0}\big[\mathrm{tr}\{A_{2ij}^2(\theta_0)\}\big] \\
=\ & p_i^2(N)p_j^2(N)o(1).
\end{aligned}$$

It follows that, for $1 \le i \ne j \le s+1$, we have

$$\frac{1}{p_i(N)p_j(N)}\frac{\partial^2 l_\theta(\tilde{y})}{\partial\theta_i\partial\theta_j}\Big|_{\theta=\theta_0} = \frac{\mathrm{tr}\{\Psi_{2ij}(\theta_0)\}}{p_i(N)p_j(N)} + o_p(1).$$

Thus, we obtain (S5.1) with $I_N(\theta_0) = [I_{Nij}(\theta_0)]_{1\le i,j\le s+1}$ and

$$I_{Nij}(\theta_0) = \frac{\mathrm{tr}\{\Psi_{2ij}(\theta_0)\}}{p_i(N)p_j(N)}.$$

Next, by Lemma 7.2 in Jiang (1996), we need to show that there are sequences of positive numbers, $q_i(N)$, such that $p_i(N)q_i(N) \to \infty$, $1 \le i \le s+1$,

$$\frac{1}{p_i(N)p_j(N)p_l(N)}\sup_{\theta\in\Theta_N}\left|\frac{\partial^3 l_\theta(\tilde{y})}{\partial\theta_i\partial\theta_j\partial\theta_l}\right| \xrightarrow{\mathrm{P}} 0,\ \ 1 \le i,j,l \le s+1, \qquad \text{(S5.2)}$$

where $\Theta_N = \{\|\theta_i - \theta_{i0}\| < q_i(N), i = 1, \cdots, s+1\}$. Deine

$$q_i(N) = \frac{\sqrt{\min_{1\le j\le s+1} p_j(N)}}{p_i(N)},\ \ i = 1, \ldots, s+1.$$

Note that we have the following expression: $\partial^3 l_\theta(\tilde{y})/\partial\theta_i\partial\theta_j\partial\theta_l =$

$$
\begin{aligned}
&\mathrm{tr}\Big\{V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_i}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_j}\Big\}\\
&+\mathrm{tr}\Big\{V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_i}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_j}\Big\}\\
&-2\tilde{y}'V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_i}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_j}V^{-1}(\Phi,\theta)\tilde{y}\\
&-2\tilde{y}'V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_i}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_j}V^{-1}(\Phi,\theta)\tilde{y}\\
&-2\tilde{y}'V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_i}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_j}V^{-1}(\Phi,\theta)\frac{\partial V(\Phi,\theta)}{\partial\theta_l}V^{-1}(\Phi,\theta)\tilde{y}\\
&\equiv \mathrm{tr}\{\Psi_{3lij}(\theta)\}+\mathrm{tr}\{\Psi_{3ilj}(\theta)\}-2W_N'(\theta_0)b(\theta_0)A_{3lij}(\theta)b'(\theta_0)W_N(\theta_0)\\
&-2W_N'(\theta_0)b(\theta_0)A_{3ilj}(\theta)b'(\theta_0)W_N(\theta_0)-2W_N'(\theta_0)b(\theta_0)A_{3ijl}(\theta)b'(\theta_0)W_N(\theta_0),
\end{aligned}
$$

$1 \le i,j,l \le s+1$. We then have $\sup_{\theta\in\Theta_N}|\partial^3 l_\theta(\tilde{y})/\partial\theta_i\partial\theta_j\partial\theta_l| \le$

$$
\begin{aligned}
&\sup_{\theta\in\Theta_N}|\mathrm{tr}\{\Psi_{3lij}(\theta)\}|+\sup_{\theta\in\Theta_N}|\mathrm{tr}\{\Psi_{3ilj}(\theta)\}|\\
&+2\left[\sup_{\theta\in\Theta_N}|W_N'(\theta_0)b(\theta_0)A_{3lij}(\theta)b'(\theta_0)W_N(\theta_0)|+\cdots\right],
\end{aligned}
$$

where $\cdots$ denotes two similar terms to the first term in the square brackets. By assumption (iv), we have $\sup_{\theta\in\Theta_N}|\mathrm{tr}\{\Psi_{3lij}(\theta)\}| = p_i(N)p_j(N)p_l(N)o(1)$ and

$$\sup_{\theta\in\Theta_N}|\mathrm{tr}\{\Psi_{3ilj}(\theta)\}| = p_i(N)p_j(N)p_l(N)o(1).$$

Also, for any $\theta\in\Theta_N$, we have $W_N'(\theta_0)b(\theta_0)A_{3lij}(\theta)b'(\theta_0)W_N(\theta_0) \le$

$$
\begin{aligned}
&\|A_{3lij}(\theta)\|W_N'(\theta_0)b(\theta_0)b'(\theta_0)W_N(\theta_0)\\
\le\ &\left\{\sup_{\theta\in\Theta_N}\|A_{3lij}(\theta)\|\right\}\mathrm{tr}\big\{b'(\theta_0)W_N(\theta_0)W_N'(\theta_0)b(\theta_0)\big\},
\end{aligned}
$$

using the fact that $A \leq \|A\| I_n$ for an $n \times n$ nonnegative definite matrix A, and $A \leq B$ implies $v'Av \leq v'Bv$ for symmetric matrices A, B and vector v. It follows that

$$\mathrm{E}\left\{\sup_{\theta\in\Theta_N} W_N'(\theta_0)b(\theta_0)A_{3lij}(\theta)b'(\theta_0)W_N(\theta_0)\right\} \leq \left\{\sup_{\theta\in\Theta_N} \|A_{3lij}(\theta)\|\right\} \mathrm{E}\left[\mathrm{tr}\big\{b'(\theta_0)b(\theta_0)\big\}\right].$$

Thus, again by assumption (iv), it is easy to verify (S5.2).

Finally, by Lemma 7.2 of Jiang (1996), we need to show boundedness in probability of $A_N(\theta_0) = [A_{N1}(\theta_0), \cdots, A_{N,s+1}(\theta_0)] =$

$$\left[\frac{1}{p_1(N)}\frac{\partial l_\theta(\tilde{y})}{\partial\theta_1}\Big|_{\theta=\theta_0}, \cdots, \frac{1}{p_{s+1}(N)}\frac{\partial l_\theta(\tilde{y})}{\partial\theta_{s+1}}\Big|_{\theta=\theta_0}\right].$$

Noting that, for $i = 1, \cdots, s+1$, we have $\partial l_\theta(\tilde{y})/\partial\theta_i|_{\theta=\theta_0} =$

$$\begin{aligned}
& \mathrm{tr}\Big\{V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_i}\Big\} - \tilde{y}'V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_i}V^{-1}(\Phi,\theta_0)\tilde{y} \\
= \;& \mathrm{tr}\Big\{V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_i}\Big\} \\
& - W_N'(\theta_0)b(\theta_0)\Phi V^{-1}(\Phi,\theta_0)\frac{\partial V(\Phi,\theta_0)}{\partial\theta_i}V^{-1}(\Phi,\theta_0)\Phi'b'(\theta_0)W_N(\theta_0) \\
= \;& \mathrm{tr}\{\Psi_{1i}(\theta_0)\} - W_N'(\theta_0)A_{1i}(\theta_0)W_N(\theta_0).
\end{aligned}$$

Thus, $A_{Ni}(\theta_0)$ can be written as $\tilde{A}_{1Ni}(\theta_0) - \tilde{A}_{2Ni}(\theta_0)$, where

$$\begin{aligned}
\tilde{A}_{1Ni}(\theta_0) &= \frac{1}{p_i(N)}\Big(\mathrm{tr}\{\Psi_{1i}(\theta_0)\} - \mathrm{E}_{\theta_0}\big[\mathrm{tr}\{A_{1i}(\theta_0)\}\big]\Big) \\
\tilde{A}_{2Ni}(\theta_0) &= \frac{1}{p_i(N)}\Big(W_N'(\theta_0)A_{1i}(\theta_0)W_N(\theta_0) - \mathrm{E}_{\theta_0}\big[\mathrm{tr}\{A_{1i}(\theta_0)\}\big]\Big).
\end{aligned}$$

By assumption (v), we have $\tilde{A}_{1Ni}(\theta_0) = O(1)$. Furthermore, by assumption (v) and properties of quadratic forms of normal random variables, we have

$\mathrm{E}_{\theta_0}\{\tilde{A}_{2Ni}(\theta_0)\} = 0$ and

$$\mathrm{Var}_{\theta_0}\{\tilde{A}_{2Ni}(\theta_0)\} = \frac{\mathrm{Var}_{\theta_0}\big[\mathrm{tr}\{A_{1i}(\theta_0)\}\big] + 2\mathrm{E}_{\theta_0}\big[\mathrm{tr}\{A_{1i}^2(\theta_0)\}\big]}{p_i^2(N)} = O(1),$$

which implies $\tilde{A}_{2Ni}(\theta_0) = O_{\mathrm{P}}(1)$. This completes the proof by Lemma 7.2 of Jiang (1996).

S6 Proof of Theorem 4

Continue with the notation introduced in the proof of Theorem 3. Also write E and Var for E_{θ_0} and Var_{θ_0}, respectively, for notation simplicity.

For any $a \in \mathrm{R}^{s+1} \setminus \{0\}$, let $c_a = (c_{a1}, \cdots, c_{a\underline{s+1}})' = \Sigma^{-1/2}(\theta_0)a$, we have

$$a'\Sigma^{-1/2}(\theta_0)A_N'(\theta_0) = c_a'A_N'(\theta_0) \equiv c_a'\tilde{A}_{1N}'(\theta_0) - c_a'\tilde{A}_{2N}'(\theta_0). \tag{S6.1}$$

Assumption (v+) implies $c_a'\tilde{A}_{1N}'(\theta_0) = o(1)$.

Next, write $\Delta_N(\theta_0) = \sum_{i=1}^{s+1} c_{ai}\{A_{1i}(\theta_0)/p_i(N)\}$. Then, we have

$$\begin{aligned} c_a'\tilde{A}_{2N}'(\theta_0) &= W_N'(\theta_0)\Delta_N(\theta_0)W_N(\theta_0) - \mathrm{E}_{\theta_0}\big\{W_N'(\theta_0)\Delta_N(\theta_0)W_N(\theta_0)\big\} \\ &= W_N'(\theta_0)\big[\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\}\big]W_N(\theta_0) \\ &\quad + W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0) - \mathrm{E}\{W_N'(\theta_0)\Delta_N(\theta_0)W_N(\theta_0)\} \\ &\equiv \Omega_1 + \Omega_2, \end{aligned}$$

with $\Omega_r, r = 1, 2$ defined in obvious ways. It is easy to show $\mathrm{E}(\Omega_1) = 0$. Also, we have

$$\mathrm{Var}(\Omega_1) = \mathrm{E}\big\{\mathrm{E}(\Omega_1^2|\tilde{Z})\big\} = \mathrm{E}\big\{\mathrm{Var}(\Omega_1|\tilde{Z}) + \mathrm{E}^2(\Omega_1|\tilde{Z})\big\}. \tag{S6.2}$$

By properties of quadratic forms of random variables, we have

$$\mathrm{E}(\Omega_1|\tilde{Z}) = \mathrm{tr}\big[\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\}\big] = \sum_{i=1}^{s+1} \frac{c_{ai}}{p_i(N)} \mathrm{tr}\big[A_{1i}(\theta_0) - \mathrm{E}\{A_{1i}(\theta_0)\}\big].$$

Thus, by assumption (v+), and the Cauchy-Schwarz inequality, we have

$$\mathrm{E}\{\mathrm{E}^2(\Omega_1|\tilde{Z})\} \leq (s+1) \sum_{i=1}^{s+1} \frac{c_{ai}^2}{p_i^2(N)} \mathrm{Var}[\mathrm{tr}\{A_{1i}(\theta_0)\}] = o(1).$$

Next, by properties of quadratic forms of normal random variables, we have

$$\begin{aligned} \mathrm{Var}(\Omega_1|\tilde{Z}) &= 2\mathrm{tr}\big([\Delta_N(\theta_0) - \mathrm{E}\{\Delta_N(\theta_0)\}]^2\big) \\ &= 2c_a' \mathrm{Tc}(A_{1i}(\theta_0)/p_i(N), 1 \leq i \leq s+1) c_a. \end{aligned}$$

Thus, by the last assumption of the theorem, we have $\mathrm{E}\{\mathrm{Var}(\Omega_1|\tilde{Z})\} =$

$$2a'\Sigma^{-1/2}(\theta_0)\mathrm{E}\{\mathrm{Tc}(A_{1i}(\theta_0)/p_i(N), 1 \leq i \leq s+1)\}\Sigma^{-1/2}(\theta_0)a = o(1).$$

Combining the above results with (S6.2), we have $\Omega_1 = o_{\mathrm{P}}(1)$.

Now we consider Ω_2. By properties of quadratic forms of normal random variables, and assumption (i) of Theorem 3, it can be shown that $\mathrm{Var}(\Omega_2) =$

$$\mathrm{Var}\big[W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0)\big] = 2\mathrm{tr}\big[\mathrm{E}^2\big\{\Delta_N(\theta_0)\big\}\big] \longrightarrow c_a'\Sigma(\theta_0)c_a = |a|^2,$$

hence $\mathrm{tr}[\mathrm{E}^2\{\Delta_N(\theta_0)\}] \to |a|^2/2$. On the other hand, by the definition of $\Sigma(\theta_0)$

[above (3.6) in MJ22], it can be shown that $\lambda_{\min}(\Sigma(\theta_0)) > 0$. It follows that

$$\lambda_{\max}\left[\mathrm{E}^2\{\Delta_N(\theta_0)\}\right] \le c|a|^2 \sum_{i=1}^{s+1} \frac{1}{p_i^2(N)}$$

for some constant $c > 0$. Combining the above results, it follows that

$$\frac{\lambda_{\max}\left[\mathrm{E}^2\{\Delta_N(\theta_0)\}\right]}{\mathrm{tr}\left[\mathrm{E}^2\{\Delta_N(\theta_0)\}\right]} \le c \sum_{i=1}^{s+1} \frac{1}{p_i^2(N)} \to 0,$$

for some constant c. Now applying Theorem 5.1 in Jiang (1996), we have

$$\frac{W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0) - \mathrm{E}\{W_N'(\theta_0)\Delta_N(\theta_0)W_N(\theta_0)\}}{\sqrt{\mathrm{Var}\left[W_N'(\theta_0)\mathrm{E}\{\Delta_N(\theta_0)\}W_N(\theta_0)\right]}} \xrightarrow{\mathrm{d}} N(0,1),$$

hence, combined with an earlier result, we have $\Omega_2 \xrightarrow{\mathrm{d}} N(0, |a|^2)$. Note that

$$\mathrm{E}\{W_N'(\theta_0)\Delta_N(\theta_0)W_N(\theta_0)\} = \mathrm{E}[W_N'(\theta_0)\mathrm{E}\{\Delta(\theta_0)\}W_N(\theta_0)]$$

by assumption (i) of Theorem 3.

Combining the results, and by the arbitrariness of a, we have

$$\Sigma^{-\frac{1}{2}}(\theta_0)A_N(\theta_0) \xrightarrow{\mathrm{d}} N(0, I_{s+1}).$$

The proof is complete by applying Lemma 7.2 of Jiang (1996).

S7 Proofs of Theorem 5 and Theorem 6 ($\hat{\beta}$ parts)

In this section, we provide the proofs of Theorem 5 and Theorem 6 for the ML estimator of β only. The proofs for the ML estimators of the variance components are similar to those for the REML estimators, and therefore omitted.

S7.1 Proof of Theorem 5 ($\hat{\beta}$ part)

By Lemma 7.2 of Jiang (1996), it suffices to show the boundedness of

$$A_{N1}(\theta_0) = \frac{1}{p_1(N)} \frac{\partial l_\theta(y)}{\partial \theta_1}\Big|_{\theta=\theta_0} = -\frac{2}{p_1(N)} X'V^{-1}(\tau_0^2, \sigma_0^2) b'(\theta_0) W_N(\theta_0).$$

By Assumptions (i) and (v) in Theorem 5, it can be shown that $\mathrm{E}\{A_{N1}(\theta_0)\} = 0$,

$$\mathrm{tr}\big[\mathrm{Var}\{A_{N1}(\theta_0)\}\big] = \frac{4}{p_1^2(N)} \mathrm{tr}\big[X'V^{-1}(\tau_0^2, \sigma_0^2)\mathrm{E}\{b'(\theta_0)b(\theta_0)\}V^{-1}(\tau_0^2, \sigma_0^2)X\big] = O(1).$$

The boundedness in probability of $|A_{N1}(\theta_0)|$ then follows. The boundedness in probability of $p_1(N)(\hat{\beta} - \beta_0)$, hence consistency of $\hat{\beta}$ then follows by Lemma 7.2 (i) of Jiang (1996).

S7.2 Proof of Theorem 6 ($\hat{\beta}$ part)

First, we establish the asymptotical normality of $A_{N1}(\theta_0)$. Note that $p_1(N) \sim \sqrt{N}$, and

$$\begin{aligned} A_{N1}(\theta_0) &= -\frac{2}{\sqrt{N}} X'V^{-1}(\tau_0^2, \sigma_0^2)\big[b'(\theta_0) - \mathrm{E}\{b'(\theta_0)\}\big] W_N(\theta_0) \\ &\quad -\frac{2}{\sqrt{N}} X'V^{-1}(\tau_0^2, \sigma_0^2)\mathrm{E}\{b'(\theta_0)\} W_N(\theta_0) \\ &\equiv \tilde{A}_{N11}(\theta_0) + \tilde{A}_{N12}(\theta_0). \end{aligned}$$

For $\tilde{A}_{N11}(\theta_0)$, we have $\mathrm{E}\{\tilde{A}_{N11}(\theta_0)\} = 0$. Furthermore, by the law of iterated expectations, assumption (i) of Theorem 5, and assumption (vi) of Theorem

6, we have

$$\begin{aligned}
& \mathrm{tr}\big[\mathrm{Var}\{\tilde{A}_{N11}(\theta_0)\}\big] \\
= \; & \frac{4}{N}\mathrm{tr}\Big(X'V^{-1}(\tau_0^2,\sigma_0^2)\big[\mathrm{E}\{b'(\theta_0)b(\theta_0)\} - \mathrm{E}\{b'(\theta_0)\}\mathrm{E}\{b(\theta_0)\}\big]V^{-1}(\tau_0^2,\sigma_0^2)X\Big) \\
= \; & o(1).
\end{aligned}$$

It follows that $\tilde{A}_{N11}(\theta_0) = o_{\mathrm{P}}(1)$.

For $\tilde{A}_{N12}(\theta_0)$, note that every component of $W_N(\theta_0)$ is distributed independently as $\mathrm{N}(0,1)$. Thus, we have $\mathrm{Var}\{\tilde{A}_{N12}(\theta_0)\} =$

$$\begin{aligned}
& \frac{4}{N}X'V^{-1}(\tau_0^2,\sigma_0^2)\mathrm{E}\{b'(\theta_0)\}\mathrm{E}\{b(\theta_0)\}V^{-1}(\tau_0^2,\sigma_0^2)X \\
= \; & \frac{4}{N}X'V^{-1}(\tau_0^2,\sigma_0^2)\big[\mathrm{E}\{b'(\theta_0)\}\mathrm{E}\{b(\theta_0)\} - V(\tau_0^2,\sigma_0^2)\big]V^{-1}(\tau_0^2,\sigma_0^2)X \\
& + \frac{4}{N}X'V^{-1}(\tau_0^2,\sigma_0^2)X;
\end{aligned}$$

hence, by assumption (vi) of Theorem 6, it follows that

$$\mathrm{Var}\{\tilde{A}_{N12}(\theta_0)\} = \frac{4}{N}X'V^{-1}(\tau_0^2,\sigma_0^2)X + o(1).$$

Thus, by the Hájek-Sidak theorem (e.g., Jiang (2022), Example 6.6), it can be shown that

$$\left\{\frac{4}{N}X'V^{-1}(\tau_0^2,\sigma_0^2)X\right\}^{-1/2}A_{N1}(\theta_0) \xrightarrow{\mathrm{d}} N(0, I_q).$$

Furthermore, by noting the following relation: $-A_{N1}(\theta_0) = I_{N1}(\theta_0)p_1(N)(\hat{\beta} - \beta_0) + o_{\mathrm{P}}(1)$, with $I_{N1}(\theta_0) = \frac{2}{N}X'V^{-1}(\tau_0^2,\sigma_0^2)X$, it follows that

$$\sqrt{N}(\hat{\beta} - \beta_0) \xrightarrow{\mathrm{d}} N\Big(0, \left\{\frac{1}{N}X'V^{-1}(\tau_0^2,\sigma_0^2)X\right\}^{-1}\Big).$$

The asymptotical normality result for $\hat{\beta}$ then follows.

S8 ML and REML: Difference and similarity

To illustrate the difference and similarity between ML and REML, let us use the balanced one-way random effects model (e.g., Jiang and Nguyen (2021)) for illustration. The model can be expressed as

$$y_{ij} \quad = \quad \mu + \alpha_i + \epsilon_{ij},$$

$i = 1, \dots, m$, $j = 1, \dots, n$, where i represents the group (e.g., subject, community) and n is the group size, that is, the number of observations in the training data that belong to group i, which is assumed to be the same for difference groups (hence explaining the term "balanced"). Furthermore, y_{ij} is the outcome of interest, μ is an unknown mean, α_i is a group-specific random effect, and ϵ_{ij} is an error. It is assumed that the random effects and errors are independent with $\alpha_i \sim N(0, \tau^2)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$, where $\sigma^2, \tau^2 > 0$ are unknown variances. The model can be expressed in a vector-matrix form:

$$y = X\beta + Z\alpha + \epsilon, \tag{S8.1}$$

where $y = (y_i)_{1 \le i \le m}$, $y_i = (y_{ij})_{1 \le j \le n_i}$, $X = 1_m \otimes 1_n$, $\alpha = (\alpha_i)_{1 \le i \le m}$, $\beta = \mu$, $Z = I_m \otimes 1_n$, $I_n, 1_n$ denote the n-dimensional identity matrix and vector of 1's, respectively, and $\otimes$ denotes the Kronecker product, $\epsilon_i = (\epsilon_{ij})_{1 \le j \le n_i}$, $\epsilon =$

$(\epsilon_i)_{1\le i\le m}$.

Standard methods of estimation for the linear mixed model (S8.1) are maximum likelihood (ML) and restricted maximum likelihood (REML). For ML, the joint pdf of y, under the LMM (S8.1), can be expressed as

$$f(y) = \frac{1}{(2\pi)^{N/2}|V|^{1/2}} \exp\big\{ -\frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta)\big\},$$

where $N = mn$ is the dimension of y and $V = ZZ'\tau^2 + I_N\sigma^2$. Thus, the log-likelihood function is given by

$$l(\beta, \tau^2, \sigma^2) = c - \frac{1}{2}\log(|V|) - \frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta), \qquad \text{(S8.2)}$$

where c is a constant. By maximizing the log-likelihood function (S8.2) with the parameters β, τ^2, σ^2, one obtains the ML estimator $\tilde{\psi} = (\tilde{\beta}, \tilde{\tau}^2, \tilde{\sigma}^2)$ of $\psi = (\beta, \tau^2, \sigma^2)$.

In general, the ML estimator of the variance components $\theta = (\tau^2, \sigma^2)$ are biased, and the bias can be severe as it may not vanish as the sample size increases, if the number of the fixed effects is proportional to the sample size (e.g., Jiang (1996)). In some cases, the fixed effects β may be viewed as nuisance parameter, while the main interest is the variance components θ. In order to estimate the variance components of main interest without having to deal with the nuisance parameters, one can apply a transformation to the data to eliminate the (nuisance) fixed effects, then use the transformed data to estimate the variance components of interest. This method is referred to as restricted maximum likelihood (REML).

Specifically, let Φ be an $N \times (N-1)$ matrix satisfying rank$(\Phi) = N-1$ and $\Phi' X = 0$. Then, define $\tilde{y} = \Phi' y$. It is easy to see that $\tilde{y} \sim N(0, \Phi' V \Phi)$. It follows that the joint pdf of $\tilde{y}$ can be expressed as

$$f_R(\tilde{y}) = \frac{1}{(2\pi)^{(N-1)/2} |\Phi' V \Phi|^{1/2}} \exp\big\{ -\frac{1}{2} \tilde{y}' (\Phi' V \Phi)^{-1} \tilde{y} \big\},$$

where the subscript R corresponds to "restricted". Thus, the log-likelihood based on $\tilde{y}$, which we call restricted log-likelihood, is given by

$$l_R(\tau^2, \sigma^2) = c - \frac{1}{2} \log(|\Phi' V \Phi|) - \frac{1}{2} \tilde{y}' (\Phi' V \Phi)^{-1} \tilde{y}, \tag{S8.3}$$

where c is a constant. By maximizing the restricted log-likelihood (S8.3) with respect to the parameters $\theta = (\tau^2, \sigma^2)$, one obtains the REML estimator $\hat{\theta}$ of θ. Note that, although the REML estimator is defined through a transforming matrix Φ, the REML estimator, in fact, does not depend on Φ. What is more, the REML estimator loses no information in estimating the parameters of variance components; in fact, asymptotically, REML has superiority over ML, if the number of fixed effects is increasing with the sample size at a sufficiently fast rate. More detail can be found in Chapter 1 of Jiang and Nguyen (2021).

Furthermore, the proofs of the asymptotic properties of ML and REML estimators are both based on a central limit theorem for quadratic forms of random variables, that is, Lemma 7.2 of Jiang (1996). Thus, starting from the (restricted) log-likelihood function, we can follow the same lines of proofs to establish standard asymptotic properties of the ML and REML estimators, including consistency and asymptotic normality. As a result, the proofs and derivations for the ML and REML of variance components are similar. The proofs for ML are simpler because the log-likelihood is somewhat easier to handle compared to the restricted log-likelihood. There are, of course, also differences in the assumptions and proofs between ML and REML, that require additional conditions (see Theorem 5 and Theorem 6).

S9 Another simulation study based on a real data example

We further illustrate the finite-sample performance with a real-data from the Nationwide Insurance Company (Nationwide). The data have been studied by Lewis, Maceachern and Lee (2021) and can be downloaded from *https:// github. com/jrlewi/brl$m$_paper/.*

Many of Nationwide's insurance policies are sold via agencies, which provide direct service to policy holders. The contractual agreements between Nationwide and these agencies vary. Identifiers such as agency/agent names are removed. Likewise, agency types (identifying the varying contractual agreements) have been de-identified to protect the proprietary nature of the data. The data are grouped by agency types, and there are $m = 10$ agency types in the analysis. We delete the agencies that closed during the 2010–2012 period. In the end, the numbers of agencies for the ten agency types are 2549, 6, 54, 51, 8, 6, 81, 8, 5 and 82, respectively.

Here, we intend to carry out a real-data based simulation not only to study the impact of the misclassification but also how the latter interact with data unbalancedness. We consider the following LMM: $y_{ij} = \beta_0 + x_{ij}'\beta + \alpha_i + e_{ij}$, where the covariate vector x_{ij} consists of the square root of the household count in 2010, and two other different size/experience measures related to the number of employees associated with the agency; the response, y_{ij} is the square root of household count in 2012; the group-specific random effects, α_i, and errors, e_{ij}, are assumed to be independent with $\alpha_i \sim N(0, \tau^2)$ and $e_{ij} \sim N(0, \sigma^2)$. Here $i = 1, \cdots 10$, $j = 1, \cdots, n_i$ with $n_1 = 2549$, $n_2 = 6$, $n_3 = 54$, $n_4 = 51$, $n_5 = 8$, $n_6 = 6$, $n_7 = 81$, $n_8 = 8$, $n_9 = 5$, and $n_{10} = 82$.

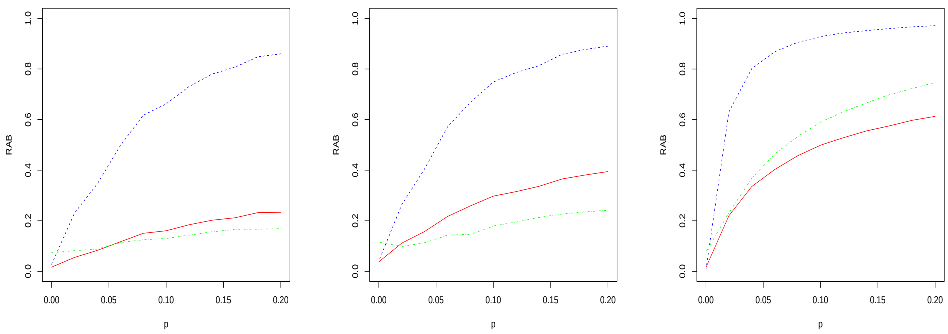
To study the influence of data unbalancedness, note that group 1 is much larger than the other groups. Thus, a subset of size $\tilde{n}_1$ is randomly chosen from group 1, which replaces group 1; the other groups are unchanged. Here $\tilde{n}_1 = 50, 200, 2000$. As the size of $\tilde{n}_1$ increases, so does the degree of unbalancedness in the data.

To study the impact of misclassification on the ML estimation, we randomly select a subset from each group, and replace the group index with other group indexes uniformly. The proportion of misclassified group indexes, p, ranges between 0 and 0.2.

For each value of p, we run 200 simulations and report the averaged relative absolute bias (RAB), defined as $\mathrm{RAB}_{\tau^2}(p) = |\hat{\tau}^2 - \hat{\tau}^2(p)|/\hat{\tau}^2$, $\mathrm{RAB}_{\sigma^2}(p) = |\hat{\sigma}^2 - \hat{\sigma}^2(p)|/\hat{\sigma}^2$, and $\mathrm{RAB}_{\beta}(p) = (1/4)\sum_{k=1}^{4}|\hat{\beta}_{k-1} - \hat{\beta}_{k-1}(p)|/|\hat{\beta}_{k-1}|$, where $\hat{\beta}_{k-1}(p), \hat{\tau}^2(p), \hat{\sigma}^2(p)$ are the MLEs of β_{k-1}, τ^2, σ^2, respectively, under the misclassification with the given p, and $\hat{\beta}_{k-1}, \hat{\tau}^2, \hat{\sigma}^2$ are the MLEs without misclassification ($k = 1, 2, 3, 4$). See Figure 1.

It can be seen from Figure 1 that the averaged relative absolute bias (RAB) increases with p for every parameter. This suggests that parameter estimation gets worse as the degree of misclassification increases. However, under different degree of data unbalancedness, every parameter has its own path of being impacted by the misclassification.

Specifically, when the degree of data unbalancedness is relatively small, that is, $\tilde{n}_1 = 50$, the magnitude of increase is the largest for τ^2; by contrast, the magnitudes of increase are very small for β and σ^2. In addition, the patterns of increase are similar for β and σ^2. When the degree of data unbalancedness is moderate, that is, $\tilde{n}_1 = 200$, the magnitude of increase is still the largest for τ^2. As for β and σ^2, although the rate of increase is still small for β, the rate of increase becomes large for σ^2. Finally, when the degree of data unbalancedness is most severe, that is, $\tilde{n}_1 = 2000$, the rate of increase becomes very fast for τ^2. Furthermore, although the increasing patterns remain similar for β and σ^2, the rate of increase for β is faster than that for σ^2.



(a) $\tilde{n}_1 = 50$ (b) $\tilde{n}_1 = 200$ (c) $\tilde{n}_1 = 2000$

Figure 1: Trend of Empirical RAB as p increases for MLE of β (dotted, green),τ^2 (dashed, blue), and σ^2 (solid, red)

From a theoretical standpoint, the convergence rates for different parameter estimation are different (Theorem 5). Specifically, the convergence rate for σ^2 and the slope coefficients of β is $O(N^{-1/2})$ with $N = \sum_{i=1}^{m} n_i$, while the convergence rate for τ^2 and the intercept β_0 is $O(m^{-1/2})$. Note that the RAB for β is defined as the average RAB for different β components; thus, the overall convergence rate for β is somewhere between $O(N^{-1/2})$ and $O(m^{-1/2})$. This may explain why the magnitude of increase in RAB for τ^2 is much larger than those for β and σ^2. It is also seen that there are some switching of orders between β and σ^2 depending on the data unbalancedness. See discussion in Section 5.

Bibliography

Jiang, J. (1996). REML estimation: Asymptotic behavior and related topics. *Ann. Statist.* **24**, 255–286.

Jiang, J. and Nguyen, T. (2021). *Linear and Generalized Linear Mixed Models and Their Applications* (2 ed.). Springer, New York.

Jiang, J. (2022). *Large Sample Techniques for Statistics* (2 ed.). Springer, New York.

Lewis, J. R. , Maceachern, S. N. , and Lee, Y. (2021). Bayesian restricted likelihood methods: conditioning on insufficient statistics in bayesian regression. *Bayesian Anal.* **16**, 1393–1462.