Economic Review (Otaru University of Commerce),
Vo1. 68, No. 2 & 3, 7-58, December, 2017.

Supplement to the paper "Expected predictive least squares for model selection in covariance structures" —Higher-order bias corrections and correlation structures

Haruhiko Ogasawara

This article supplements Ogasawara (2017).

S1. Higher-order bias corrections for least squares

S1.1 $\mathrm{ALS}_{\mathrm{NTG}}$, $\mathrm{TLS}_{\mathrm{NTG}}$, and $\mathrm{CALS}_{\mathrm{NTG}}$ when $\hat{\mathbf{W}}_{\mathbf{s}} = n\,\widehat{\mathrm{acov}_{\mathrm{NT}}(\mathbf{s})}$ by NT-GLS for covariance structures

S1.1 1 Preliminary results

$$\mathrm{E}_g^{(\mathbf{s})}(\mathbf{S}) = \mathbf{\Sigma}_{\mathrm{T}},\ n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s}) = 2\mathbf{D}_p^{+}(\mathbf{\Sigma}_{\mathrm{T}} \otimes \mathbf{\Sigma}_{\mathrm{T}})\mathbf{D}_p^{+}{}',$$

$$\mathbf{D}_p^{+} = (\mathbf{D}_p{}'\mathbf{D}_p)^{-1}\mathbf{D}_p{}',\ \mathrm{vec}(\mathbf{S}) = \mathbf{D}_p \mathrm{v}(\mathbf{S}) = \mathbf{D}_p\mathbf{s},$$

$$\mathbf{N}_p = \mathbf{D}_p\mathbf{D}_p^{+} = \mathbf{D}_p^{+}{}'\mathbf{D}_p{}'$$ (symmetrizer; Holmquist, 1988, p.275; Kano, 1997, p.182; Magnus & Neudecker, 1999, p.46),

$$\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}^{-1} = (1/2)\mathbf{D}_p{}'(\mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \mathbf{\Sigma}_{\mathrm{T}}^{-1})\mathbf{D}_p.$$

The last result is confirmed by

$$\{2\mathbf{D}_p^{+}(\mathbf{\Sigma}_{\mathrm{T}} \otimes \mathbf{\Sigma}_{\mathrm{T}})\mathbf{D}_p^{+}{}'\}\{(1/2)\mathbf{D}_p{}'(\mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \mathbf{\Sigma}_{\mathrm{T}}^{-1})\mathbf{D}_p\}$$
$$= \mathbf{D}_p^{+}(\mathbf{\Sigma}_{\mathrm{T}} \otimes \mathbf{\Sigma}_{\mathrm{T}})\mathbf{N}_p(\mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \mathbf{\Sigma}_{\mathrm{T}}^{-1})\mathbf{D}_p$$
$$= \mathbf{D}_p^{+}\mathbf{N}_p(\mathbf{\Sigma}_{\mathrm{T}} \otimes \mathbf{\Sigma}_{\mathrm{T}})(\mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \mathbf{\Sigma}_{\mathrm{T}}^{-1})\mathbf{D}_p$$

$$= \mathbf{D}_p^{+}\mathbf{N}_p\mathbf{D}_p = (\mathbf{D}_p'\mathbf{D}_p)^{-1}\mathbf{D}_p'\mathbf{D}_p(\mathbf{D}_p'\mathbf{D}_p)^{-1}\mathbf{D}_p'\mathbf{D}_p$$
$$= \mathbf{I}_{(p^*)},$$

where $(\boldsymbol{\Sigma}_{\mathrm{T}} \otimes \boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{N}_p = \mathbf{N}_p(\boldsymbol{\Sigma}_{\mathrm{T}} \otimes \boldsymbol{\Sigma}_{\mathrm{T}})$ is used.

$$(\mathbf{s}-\boldsymbol{\sigma})'\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}^{-1}(\mathbf{s}-\boldsymbol{\sigma})$$
$$= (\mathbf{s}-\boldsymbol{\sigma})'(1/2)\mathbf{D}_p'(\boldsymbol{\Sigma}_{\mathrm{T}}^{-1} \otimes \boldsymbol{\Sigma}_{\mathrm{T}}^{-1})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma})$$
$$= (1/2)\mathrm{vec}'(\mathbf{S}-\boldsymbol{\Sigma})(\boldsymbol{\Sigma}_{\mathrm{T}}^{-1} \otimes \boldsymbol{\Sigma}_{\mathrm{T}}^{-1})\mathrm{vec}(\mathbf{S}-\boldsymbol{\Sigma})$$
$$= (1/2)\mathrm{tr}[\{\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}(\mathbf{S}-\boldsymbol{\Sigma})\}^2].$$

Define F as the NT-GLS discrepancy function of $\mathbf{s}$ and $\boldsymbol{\sigma}$, then

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + \frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}}) + \frac{1}{2}\frac{\partial^2 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<2>}}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>} + \frac{1}{6}\frac{\partial^3 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<3>}}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<3>}$$
$$+ O_p(n^{-2}),$$

$$\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma}_0 + \frac{\partial \boldsymbol{\sigma}_0}{\partial \boldsymbol{\theta}_0'}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) + \frac{1}{2}\frac{\partial^2 \boldsymbol{\sigma}_0}{(\partial \boldsymbol{\theta}_0')^{<2>}}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)^{<2>} + \frac{1}{6}\frac{\partial^3 \boldsymbol{\sigma}_0}{(\partial \boldsymbol{\theta}_0')^{<3>}}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)^{<3>}$$
$$+ O_p(n^{-2}),$$

$$\frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'} = -\left(\frac{\partial^2 F}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0,\,\mathbf{s}=\boldsymbol{\sigma}_{\mathrm{T}}}\right)^{-1}\frac{\partial^2 F}{\partial \boldsymbol{\theta}\,\partial \mathbf{s}'}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0,\,\mathbf{s}=\boldsymbol{\sigma}_{\mathrm{T}}}$$
$$= -\left\{2\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0 + 2\sum_{a \ge b}(\boldsymbol{\sigma}_0-\boldsymbol{\sigma}_{\mathrm{T}})'(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\cdot ab}\frac{\partial^2 \sigma_{0ab}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0'}\right\}^{-1}\left(-2\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\right)$$
$$= \left\{\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0 + \sum_{a \ge b}(\boldsymbol{\sigma}_0-\boldsymbol{\sigma}_{\mathrm{T}})'(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\cdot ab}\frac{\partial^2 \sigma_{0ab}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0'}\right\}^{-1}\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}$$
$$\equiv (\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0 + \mathbf{A}_0)^{-1}\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1},$$

where $\mathbf{x}^{<k>} = \mathbf{x} \otimes \cdots \otimes \mathbf{x}$ (k times of $\mathbf{x}$); $\dfrac{\partial^2 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<2>}}$ and $\dfrac{\partial^3 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<3>}}$ are also given by the formulas of partial derivatives in implicit functions (Ogasawara, 2007, Equations (17) and (19); 2009, Equation (3.16)). Note that when

$\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathbf{A}_0 = \mathbf{O}$ (a zero matrix).

$$\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma}_0 + \boldsymbol{\Delta}_0(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \frac{1}{2}\boldsymbol{\Delta}_0^{(2)}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^{<2>} + \frac{1}{6}\boldsymbol{\Delta}_0^{(3)}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^{<3>} + O_p(n^{-2})$$

$$= \boldsymbol{\sigma}_0 + \left\{\boldsymbol{\Delta}_0 \frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\right\}_{O_p(n^{-1/2})} + \left[\frac{1}{2}\boldsymbol{\Delta}_0 \frac{\partial^2 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<2>}}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>}\right.$$

$$\left. + \frac{1}{2}\boldsymbol{\Delta}_0^{(2)}\left\{\frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\right\}^{<2>}\right]_{O_p(n^{-1})} + \underset{(\mathrm{A})}{\Bigg[}\ \frac{1}{6}\boldsymbol{\Delta}_0 \frac{\partial^3 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<3>}}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<3>}$$

$$+ \frac{1}{2}\boldsymbol{\Delta}_0^{(2)}\left[\left\{\frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\right\} \otimes \left\{\frac{\partial^2 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<2>}}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>}\right\}\right]$$

$$\left. + \frac{1}{6}\boldsymbol{\Delta}_0^{(3)}\left\{\frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\right\}^{<3>}\ \right]_{(\mathrm{A})\,O_p(n^{-3/2})} + O_p(n^{-2})$$

$$\equiv \boldsymbol{\sigma}_0 + \boldsymbol{\Lambda}_0^{(1)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}}) + \boldsymbol{\Lambda}_0^{(2)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>} + \boldsymbol{\Lambda}_0^{(3)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<3>} + O_p(n^{-2}),$$

$$\boldsymbol{\Lambda}_0^{(1)} = \boldsymbol{\Delta}_0 \frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'},\ \ \boldsymbol{\Lambda}_0^{(2)} = \frac{1}{2}\boldsymbol{\Delta}_0 \frac{\partial^2 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<2>}} + \frac{1}{2}\boldsymbol{\Delta}_0^{(2)}\left(\frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'}\right)^{<2>},$$

$$\boldsymbol{\Lambda}_0^{(3)} = \frac{1}{6}\boldsymbol{\Delta}_0 \frac{\partial^3 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<3>}} + \frac{1}{2}\boldsymbol{\Delta}_0^{(2)}\left\{\frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'} \otimes \frac{\partial^2 \boldsymbol{\theta}_0}{(\partial \boldsymbol{\sigma}_{\mathrm{T}}')^{<2>}}\right\} + \frac{1}{6}\boldsymbol{\Delta}_0^{(3)}\left(\frac{\partial \boldsymbol{\theta}_0}{\partial \boldsymbol{\sigma}_{\mathrm{T}}'}\right)^{<3>},$$

where $\underset{(\mathrm{A})}{[}\ \cdot\ \underset{(\mathrm{A})}{]}$ is for ease of finding correspondence.

$$\mathrm{LS}_{\mathrm{NTG}} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\hat{\mathbf{W}}_{\mathrm{NT},\mathbf{s}}^{-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}),$$

$\hat{\boldsymbol{\theta}}_{\mathrm{NGLS}}$ in $\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}} = \boldsymbol{\sigma}(\hat{\boldsymbol{\theta}}_{\mathrm{NGLS}})$ minimizes $\mathrm{LS}_{\mathrm{NTG}}$,

$$\hat{\mathbf{W}}_{\mathrm{NT},\mathbf{s}} = 2\mathbf{D}_p^+(\mathbf{S} \otimes \mathbf{S})\mathbf{D}_p^{+}{}',$$

$$\mathrm{EPLS}_{\mathrm{NTG}} = \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\mathbf{W}_{\mathrm{NT}}^{-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\}.$$

S1.1.2 Bias of $\mathrm{LS}_{\mathrm{NTG}}$

$$\begin{aligned}
&\mathrm{E}_g^{(\mathbf{s})}(\mathrm{LS}_{\mathrm{NTG}})-\mathrm{EPLS}_{\mathrm{NTG}}\\
&=\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\}\\
&\quad+\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\}\\
&\quad-\mathrm{E}_g^{(\mathbf{t})}(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})\}\\
&\quad-\mathrm{E}_g^{(\mathbf{s})}\{(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_{\mathrm{T}})\}\\
&\quad(2\mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\boldsymbol{\sigma}_{\mathrm{T}}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\}=0\ \text{is used})\\
&=-2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_{\mathrm{T}})\}\\
&\quad+\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\},
\end{aligned}\qquad(\mathrm{s}1.1.1)$$

where $\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}=(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)})^{-1}$ and $\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}$ ($\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)}$) is synonymously used with $\mathbf{W}_{\mathrm{NT}}(\hat{\mathbf{W}}_{\mathrm{NT},\,\mathbf{s}})$.

The first term on the right-hand side of the last equation of (s1.1.1) is

$$\begin{aligned}
&-2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_{\mathrm{T}})\}\\
&=-2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)\}\\
&\quad(\boldsymbol{\sigma}_{\mathrm{T}}\ \text{has been validly replaced by}\ \boldsymbol{\sigma}_0)\\
&=-2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\}]_{\to O(n^{-1})+O(n^{-2})}\\
&\quad-2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\}]_{\to O(n^{-2})}\\
&\quad-2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<3>}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\}]_{\to O(n^{-2})}+O(n^{-3})\\
&=-\{n^{-1}2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})\}_{O(n^{-1})}\\
&\quad+\underset{(\mathrm{A})}{[}\,n^{-2}2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)})-n^{-2}2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)})\\
&\qquad-n^{-2}6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}]\,\underset{(\mathrm{A})}{]}_{O(n^{-2})}+O(n^{-3}),
\end{aligned}\qquad(\mathrm{s}1.1.2)$$

where $\boldsymbol{\Gamma}_0^{(2)}=\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}$ under normality;
$\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\}]=n^{-1}\boldsymbol{\Gamma}_0^{(2)}-n^{-2}\mathbf{K}_{(4)}+O(n^{-3})$ under arbitrary distributions; $\mathbf{K}_{(4)}$ is the $p^*\times p^*$ matrix of the multivariate fourth

cumulants, whose element $(\mathbf{K}_{(4)})_{ab,cd}$ corresponds to that of observable variables X_a, X_b, X_c and $X_d\,(p \ge a \ge b \ge 1;\ p \ge c \ge d \ge 1)$,

$$\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\}] = n^{-2}\boldsymbol{\Gamma}_0^{(3)} + O(n^{-3}),$$

$$(\boldsymbol{\Gamma}_0^{(3)})_{(ab,\,cd;\,ef)} = n^2\mathrm{E}_g^{(\mathbf{s})}\{(s_{ab}-\sigma_{\mathrm{T}ab})(s_{cd}-\sigma_{\mathrm{T}cd})(s_{ef}-\sigma_{\mathrm{T}ef})\} + O(n^{-1})$$

$$= \sigma_{\mathrm{T}abcdef} - \sum^{3}\sigma_{\mathrm{T}abcd}\sigma_{\mathrm{T}ef} - \sum^{6}\sigma_{\mathrm{T}acd}\sigma_{\mathrm{T}bef} + 2\sigma_{\mathrm{T}ab}\sigma_{\mathrm{T}cd}\sigma_{\mathrm{T}ef} + O(n^{-1}),$$

$\sigma_{\mathrm{T}ab\cdots f}$ is the multivariate central moment of X_a, X_b,..., X_f $(p \ge a \ge b \ge 1;\ p \ge c \ge d \ge 1;\ p \ge e \ge f \ge 1;$ Ogasawara, 2006, Equation (3.13); 2007, Lemma 1).

Under normality, the term of order $O(n^{-1})$ in (s1.1.2) is

$$-n^{-1}2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})$$

$$= -n^{-1}2\mathrm{tr}\{\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0(\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0+\mathbf{A}_0)^{-1}\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\}$$

$$= -n^{-1}2\mathrm{tr}\{(\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0+\mathbf{A}_0)^{-1}\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0\},$$

which becomes $-n^{-1}2q$ when $\mathbf{A}_0 = \mathbf{O}$.

The second term on the right-hand side of the last equation of (s1.1.1) is

$$\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\}$$

$$= \mathrm{E}_g^{(\mathbf{s})}\{(1/2)\mathrm{vec}'(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})(\mathbf{S}^{-1}\otimes\mathbf{S}^{-1}-\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\otimes\boldsymbol{\Sigma}_{\mathrm{T}}^{-1})\mathrm{vec}(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})\}. \tag{s1.1.3}$$

Let $\mathbf{M}_{\mathbf{s}} = \mathbf{S}-\boldsymbol{\Sigma}_{\mathrm{T}}$. Then,

$$\mathbf{S}^{-1} = \boldsymbol{\Sigma}_{\mathrm{T}}^{-1} - \boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1} + \boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1} - \boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}$$
$$+ \boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1} + O_p(n^{-5/2}),$$

$$\mathbf{S}^{-1}\otimes\mathbf{S}^{-1} - \boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\otimes\boldsymbol{\Sigma}_{\mathrm{T}}^{-1} = \{-\sum\nolimits_{\mathrm{sym}}^{2}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\otimes(\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1})\}_{O_p(n^{-1/2})}$$

$$+\{(\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1})^{<2>} + \sum\nolimits_{\mathrm{sym}}^{2}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\otimes(\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\mathrm{T}}^{-1})\}_{O_p(n^{-1})}$$

$$
\begin{aligned}
&+\{-\textstyle\sum_{\rm sym}^{2}\boldsymbol{\Sigma}_{\rm T}^{-1}\otimes(\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1})\\
&\quad-\textstyle\sum_{\rm sym}^{2}(\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1})\otimes(\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1})\}_{O_p(n^{-3/2})}\\
&+\{\textstyle\sum_{\rm sym}^{2}\boldsymbol{\Sigma}_{\rm T}^{-1}\otimes(\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1})\\
&\quad+(\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1})^{<2>}\\
&\quad+\textstyle\sum_{\rm sym}^{2}(\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1})\otimes(\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1}\mathbf{M}_{\mathbf{s}}\boldsymbol{\Sigma}_{\rm T}^{-1})\}_{O_p(n^{-2})}+O_p(n^{-5/2})\\
&\equiv(\mathbf{M}^{(1)})_{O_p(n^{-1/2})}+(\mathbf{M}^{(2)})_{O_p(n^{-1})}+(\mathbf{M}^{(3)})_{O_p(n^{-3/2})}+(\mathbf{M}^{(4)})_{O_p(n^{-2})}+O_p(n^{-5/2}),
\end{aligned}
$$

where $\sum_{\rm sym}^{2}\mathbf{X}=\mathbf{X}+\mathbf{X}'$.

The right-hand side of (s1.1.3) becomes

$$
\begin{aligned}
&{\rm E}_g^{(\mathbf{s})}\{(1/2){\rm vec}'(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\rm NGLS})(\mathbf{S}^{-1}\otimes\mathbf{S}^{-1}-\boldsymbol{\Sigma}_{\rm T}^{-1}\otimes\boldsymbol{\Sigma}_{\rm T}^{-1}){\rm vec}(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\rm NGLS})\}\\
&={\rm E}_g^{(\mathbf{s})}[(1/2){\rm vec}'\{\mathbf{S}-\boldsymbol{\Sigma}_{\rm T}-(\hat{\boldsymbol{\Sigma}}_{\rm NGLS}-\boldsymbol{\Sigma}_0)+\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0\}\\
&\qquad\times(\mathbf{M}^{(1)}+\mathbf{M}^{(2)}+\mathbf{M}^{(3)}+\mathbf{M}^{(4)})\\
&\qquad\times{\rm vec}\{\mathbf{S}-\boldsymbol{\Sigma}_{\rm T}-(\hat{\boldsymbol{\Sigma}}_{\rm NGLS}-\boldsymbol{\Sigma}_0)+\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0\}]+O(n^{-3})
\end{aligned}
\qquad\text{(s1.1.4)}
$$

(note that $\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0=O(1)$).

The term of order $O(n^{-1})$ in (s1.1.4) is

$$
\begin{aligned}
&n^{-1}[(1/2){\rm vec}'(\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0)n{\rm E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}{\rm vec}(\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0)\\
&\quad+{\rm vec}'(\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0)n{\rm E}_g^{(\mathbf{s})}\{\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})\}_{\to O(n^{-1})}],
\end{aligned}
\qquad\text{(s1.1.5)}
$$

which becomes zero when $\boldsymbol{\Sigma}_{\rm T}=\boldsymbol{\Sigma}_0$.

The term of order $O(n^{-2})$ in (s1.1.4) is

$$
\begin{aligned}
&n^{-2}\underset{\rm (A)}{[}\,(1/2){\rm vec}'(\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0)n^2{\rm E}_g^{(\mathbf{s})}\{\mathbf{M}^{(2)}-{\rm E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\\
&\qquad\qquad+\mathbf{M}^{(3)}+\mathbf{M}^{(4)}\}_{\to O(n^{-2})}{\rm vec}(\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0)\\
&\quad+{\rm vec}'(\boldsymbol{\Sigma}_{\rm T}-\boldsymbol{\Sigma}_0)n^2{\rm E}_g^{(\mathbf{s})}\{(\mathbf{M}^{(2)}+\mathbf{M}^{(3)})\\
&\qquad\qquad\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})\}_{\to O(n^{-2})}
\end{aligned}
\qquad\text{(s1.1.6)}
$$

$$-\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}+\mathbf{M}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<3>}$$
$$+\mathbf{M}^{(2)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}\}_{\to O(n^{-2})}$$
$$+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]}.$$

When $\boldsymbol{\sigma}_{\mathrm{T}}=\boldsymbol{\sigma}_0$, (s1.1.6) becomes

$$n^{-2}\underset{(\mathrm{A})}{[}(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]}.$$

Then,

$$\mathrm{E}_g^{(\mathbf{s})}(\mathrm{LS}_{\mathrm{NTG}})-\mathrm{EPLS}_{\mathrm{NTG}}$$
$$=n^{-1}[-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})$$
$$+(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)$$
$$+\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-1})}] \quad \text{(s1.1.7)}$$
$$+n^{-2}\underset{(\mathrm{A})}{[}2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)})-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)})$$
$$-6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}]$$
$$+(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}^{(2)}-\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}$$
$$+\mathbf{M}^{(3)}+\mathbf{M}^{(4)}\}_{\to O(n^{-2})}\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)$$
$$+\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}^{(2)}+\mathbf{M}^{(3)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-\mathrm{vec}'(\mathbf{\Sigma}_{\mathrm{T}}-\mathbf{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}^{(1)}\mathbf{D}_p\mathbf{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}+\mathbf{M}^{(1)}\mathbf{D}_p\mathbf{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<3>}$$
$$+\mathbf{M}^{(2)}\mathbf{D}_p\mathbf{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}\}_{\to O(n^{-2})}$$
$$+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]}$$
$$+O(n^{-3}).$$

(i) Under normality and $\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0=O(1)$, the term of order $O(n^{-1})$ in (s1.1.7) is $-2\mathrm{tr}(\mathbf{\Lambda}_0^{(1)})=-2\mathrm{tr}\{(\mathbf{\Delta}_0{}'\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Delta}_0+\mathbf{A}_0)^{-1}\mathbf{\Delta}_0{}'\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Delta}_0\}\neq-2q$, and the first term in $\underset{(\mathrm{A})}{[}\cdot\underset{(\mathrm{A})}{]}$ for the term of order $O(n^{-2})$ becomes $2\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Lambda}_0^{(1)}\mathbf{K}_{(4)})=0$.

(ii) Under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}}=\boldsymbol{\sigma}_0$

$$\mathrm{E}_g^{(\mathbf{s})}(\mathrm{LS}_{\mathrm{NTG}})-\mathrm{EPLS}_{\mathrm{NTG}}$$
$$=n^{-1}\{-2\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_0^{(2)})\}$$
$$+n^{-2}\underset{(\mathrm{A})}{[}\,2\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Lambda}_0^{(1)}\mathbf{K}_{(4)})-2\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Lambda}_0^{(2)}\mathbf{\Gamma}_0^{(3)})$$
$$-6\mathrm{tr}[\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Lambda}_0^{(3)}\{\mathrm{vec}(\mathbf{\Gamma}_0^{(2)})\otimes\mathbf{\Gamma}_0^{(2)}\}]$$
$$+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]}$$
$$+O(n^{-3}).$$

(iii) Under normality and $\boldsymbol{\sigma}_{\mathrm{T}}=\boldsymbol{\sigma}_0$

$$
\begin{aligned}
&\mathrm{E}_f^{(\mathbf{s})}(\mathrm{LS}_{\mathrm{NTG}}) - \mathrm{EPLS}_{\mathrm{NTG}} \\
&= n^{-1}(-2q) \\
&+ n^{-2} \underset{(\mathrm{A})}{[} -2\mathrm{tr}(\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Lambda}_0^{(2)}\mathbf{\Gamma}_{\mathrm{NT}}^{(3)}) - 6\mathrm{tr}[\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}\mathbf{\Lambda}_0^{(3)}\{\mathrm{vec}(\mathbf{\Gamma}_{\mathrm{NT}}^{(2)}) \otimes \mathbf{\Gamma}_{\mathrm{NT}}^{(2)}\}] \\
&+ (1/2)n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)} + \mathbf{M}^{(2)}) \\
&\qquad\qquad \times (\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&- n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \underset{(\mathrm{A})}{]} \\
&+ O(n^{-3}).
\end{aligned}
$$

S1.1.3 Bias correction of $\mathrm{LS}_{\mathrm{NTG}}$

Recall that $\mathrm{LS}_{\mathrm{NTG}} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})$. Define

$$
\begin{aligned}
&\mathrm{ALS}_{\mathrm{NTG}} \equiv \mathrm{LS}_{\mathrm{NTG}} + n^{-1}2q, \\
&\mathrm{TLS}_{\mathrm{NTG}} \equiv \mathrm{LS}_{\mathrm{NTG}} + n^{-1}2\mathrm{tr}(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(1)}\hat{\mathbf{\Gamma}}^{(2)}) \\
&\qquad\qquad = \mathrm{LS}_{\mathrm{NTG}} + n^{-1}2\mathrm{tr}\{(\hat{\mathbf{\Delta}}'\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\mathbf{\Delta}})^{-1}\hat{\mathbf{\Delta}}'\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\mathbf{\Gamma}}^{(2)}\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\mathbf{\Delta}}\}
\end{aligned}
$$

with $(\hat{\mathbf{\Gamma}}^{(2)})_{ab,\,cd} = s_{abcd} - s_{ab}s_{cd}\ (p \ge a \ge b \ge 1;\ p \ge c \ge d \ge 1)$, and

$$
\begin{aligned}
&\mathrm{CALS}_{\mathrm{NTG}} \equiv \mathrm{LS}_{\mathrm{NTG}} + n^{-1}2q \\
&- n^{-2} \underset{(\mathrm{A})}{[} -2\mathrm{tr}(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(2)}\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(3)}) - 6\mathrm{tr}[\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(3)}\{\mathrm{vec}(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)}) \otimes \hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(2)}\}] \\
&+ (1/2)n^2 \widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)} + \mathbf{M}^{(2)}) \\
&\qquad\qquad \times (\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&- n^2 \widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \underset{(\mathrm{A})}{]},
\end{aligned}
$$

where $\widehat{\mathrm{E}_f^{(\mathbf{s})}}\{\cdot\} = \widehat{\mathrm{E}_f^{(\mathbf{s})}\{\cdot\}}$.

The bias corrections in $\mathrm{ALS}_{\mathrm{NTG}}$, $\mathrm{TLS}_{\mathrm{NTG}}$ and $\mathrm{CALS}_{\mathrm{NTG}}$ are valid only when a structural model is true i.e., $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$.

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathrm{E}_f(\mathrm{ALS}_{\mathrm{NTG}}) - \mathrm{EPLS}_{\mathrm{NTG}} = O(n^{-2})$ and $\mathrm{E}_f(\mathrm{CALS}_{\mathrm{NTG}}) - \mathrm{EPLS}_{\mathrm{NTG}} = O(n^{-3})$.

Under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathrm{E}_g(\mathrm{TLS}_{\mathrm{NTG}}) - \mathrm{EPLS}_{\mathrm{NTG}} = O(n^{-2})$.

S1.1.4 The results for the saturated model under normality

For the saturated model with $\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}} = \mathbf{s}$, $\mathrm{E}_f(\mathrm{LS}_{\mathrm{NTG}}) = \mathrm{LS}_{\mathrm{NTG}} = 0$. Then, under normality,

$$\begin{aligned}
&\mathrm{E}_f^{(\mathbf{s})}(\mathrm{LS}_{\mathrm{NTG}}) - \mathrm{EPLS}_{\mathrm{NTG}} \\
&= -\mathrm{EPLS}_{\mathrm{NTG}} = -\mathrm{E}_f^{(\mathbf{t})}\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{t}-\mathbf{s})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{t}-\mathbf{s})\} \\
&= -\mathrm{E}_f^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})\} - \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\} \qquad \text{(s1.1.8)} \\
&= -n^{-1}2q,
\end{aligned}$$

which is an exact result. Alternatively, from an intermediate result of (s1.1.2) using $\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}} = \mathbf{s} - \mathbf{s} = \mathbf{0}$, we have

$$\begin{aligned}
&-2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)\} \\
&= -2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{s}-\boldsymbol{\sigma}_0)\} = -n^{-1}2q.
\end{aligned}$$

The corresponding result based on cross-validation by Browne and Cudeck (1989, Equation (7)) is

$$\begin{aligned}
&\mathrm{E}_f^{(\mathbf{t})}\mathrm{E}_f^{(\mathbf{s})} \underset{(\mathrm{A})}{[} (1/2)\mathrm{tr}[\{\mathbf{S}^{-1}(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})\}^2] \\
&\qquad - (1/2)\mathrm{tr}[\{\mathbf{T}^{-1}(\mathbf{T}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})\}^2] \underset{(\mathrm{A})}{]} \\
&= -\mathrm{E}_f^{(\mathbf{t})}\mathrm{E}_f^{(\mathbf{s})}[(1/2)\mathrm{tr}\{(\mathbf{I}_{(p)} - \mathbf{T}^{-1}\mathbf{S})^2\}] \qquad \text{(s1.1.9)} \\
&= -\frac{2qn^2 + 2pn - \{p^2(p+1)(p+3)/2\}}{(n-p)(n-p-1)(n-p-3)},
\end{aligned}$$

where $\mathbf{S} = \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}}$. The values of (s1.1.8) and (s1.1.9) are the same up to order $O(n^{-1})$ while the absolute value of (s1.1.9) is larger than that of (s1.1.8) when n is sufficiently large.

S1.2 $\mathrm{ALS}_{\mathrm{NTG*}}$, $\mathrm{TLS}_{\mathrm{NTG*}}$, and $\mathrm{CALS}_{\mathrm{NTG*}}$ by NT-GLS* when

$\hat{\mathbf{W}}_{\mathbf{s}} = \hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})}(\ (\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})})_{ab,\,cd} = \hat{\sigma}_{\mathrm{NGLS*},\,ac}\hat{\sigma}_{\mathrm{NGLS*},\,bd}$
$+\hat{\sigma}_{\mathrm{NGLS*},\,ad}\hat{\sigma}_{\mathrm{NGLS*},\,bc}; p \ge a \ge b \ge 1;\ p \ge c \ge d \ge 1)$ **for covariance structures**

S1.2.1 Definition

$$\begin{aligned}\mathrm{LS}_{\mathrm{NTG*}} &\equiv (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS*}})'\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS*}})\\ &= (1/2)\mathrm{vec}'(\mathbf{S} - \hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}})(\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}}^{-1} \otimes \hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}}^{-1})\mathrm{vec}(\mathbf{S} - \hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}})\\ &= (1/2)\mathrm{tr}[\{\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}}^{-1}(\mathbf{S} - \hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}})\}^2].\end{aligned}$$

S1.2.2 Bias of $\mathrm{LS}_{\mathrm{NTG*}}$ under possible non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$

The case $\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0 = O(1)$ is not dealt with in this subsection. Define $\mathrm{ELS}_{\mathrm{NTG*}} = \mathrm{E}_g^{(\mathbf{s})}(\mathrm{LS}_{\mathrm{NTG*}})$. Then,

$$\begin{aligned}&\mathrm{ELS}_{\mathrm{NTG*}} - \mathrm{EPLS}_{\mathrm{NTG}}\\ &= -2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS*}} - \boldsymbol{\sigma}_{\mathrm{T}})\}\\ &\quad +\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS*}})'(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1} - \mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1})(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS*}})\},\end{aligned} \tag{s1.2.1}$$

where $\mathrm{EPLS}_{\mathrm{NTG}}$ is as before and the first term on the right-hand side of (s1.2.1) is given as in (s1.1.2). The second term on the right-hand side of (s1.2.1) is

$$\begin{aligned}&\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS*}})'(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1} - \mathbf{\Gamma}_{\mathrm{NT}}^{(2)-1})(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS*}})\}\\ &= \mathrm{E}_g^{(\mathbf{s})}[(1/2)\mathrm{vec}'\{\mathbf{S} - \mathbf{\Sigma}_{\mathrm{T}} - (\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}} - \mathbf{\Sigma}_0)\}\\ &\quad \times(\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}}^{-1} \otimes \hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}}^{-1} - \mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \mathbf{\Sigma}_{\mathrm{T}}^{-1})\mathrm{vec}\{\mathbf{S} - \mathbf{\Sigma}_{\mathrm{T}} - (\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}} - \mathbf{\Sigma}_0)\},\end{aligned} \tag{s1.2.2}$$

where

$$\begin{aligned}&\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}}^{-1} \otimes \hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}}^{-1} - \mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \mathbf{\Sigma}_{\mathrm{T}}^{-1}\\ &= -\sum\nolimits_{\mathrm{sym}}^{2} \mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \{\mathbf{\Sigma}_{\mathrm{T}}^{-1}(\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}} - \mathbf{\Sigma}_{\mathrm{T}})\mathbf{\Sigma}_{\mathrm{T}}^{-1}\}\\ &\quad + \sum\nolimits_{\mathrm{sym}}^{2} \mathbf{\Sigma}_{\mathrm{T}}^{-1} \otimes \{\mathbf{\Sigma}_{\mathrm{T}}^{-1}(\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}} - \mathbf{\Sigma}_{\mathrm{T}})\mathbf{\Sigma}_{\mathrm{T}}^{-1}(\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}} - \mathbf{\Sigma}_{\mathrm{T}})\mathbf{\Sigma}_{\mathrm{T}}^{-1}\}\\ &\quad + \{\mathbf{\Sigma}_{\mathrm{T}}^{-1}(\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}} - \mathbf{\Sigma}_{\mathrm{T}})\mathbf{\Sigma}_{\mathrm{T}}^{-1}\} \otimes \mathbf{\Sigma}_{\mathrm{T}}^{-1}(\hat{\mathbf{\Sigma}}_{\mathrm{NGLS*}} - \mathbf{\Sigma}_{\mathrm{T}})\mathbf{\Sigma}_{\mathrm{T}}^{-1}\} + O_p(n^{-3/2})\end{aligned}$$

$$= -\sum\nolimits_{\text{sym}}^{2} \boldsymbol{\Sigma}_{\text{T}}^{-1} \otimes \{\sum\nolimits_{a,b=1}^{p} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{\cdot a} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{b\cdot} (\boldsymbol{\Lambda}_0^{(1)})_{ab\cdot} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})\}$$

$$+ \underset{(\text{A})}{[} -\sum\nolimits_{\text{sym}}^{2} \boldsymbol{\Sigma}_{\text{T}}^{-1} \otimes \{\sum\nolimits_{a,b=1}^{p} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{\cdot a} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{b\cdot} (\boldsymbol{\Lambda}_0^{(2)})_{ab\cdot} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})^{<2>}\}$$

$$+ \sum\nolimits_{\text{sym}}^{2} \boldsymbol{\Sigma}_{\text{T}}^{-1} \otimes \{\sum\nolimits_{a,b=1}^{p} \sum\nolimits_{c,d=1}^{p} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{\cdot a} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{bc} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{d\cdot}$$

$$\times (\boldsymbol{\Lambda}_0^{(1)})_{ab\cdot} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}}) (\boldsymbol{\Lambda}_0^{(1)})_{cd\cdot} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})\}$$

$$+ \{\sum\nolimits_{a,b=1}^{p} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{\cdot a} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{b\cdot} (\boldsymbol{\Lambda}_0^{(1)})_{ab\cdot} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})\}$$

$$\otimes \{\sum\nolimits_{c,d=1}^{p} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{\cdot c} (\boldsymbol{\Sigma}_{\text{T}}^{-1})_{d\cdot} (\boldsymbol{\Lambda}_0^{(1)})_{cd\cdot} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})\} \underset{(\text{A})}{]} + O_p(n^{-3/2})$$

$$\equiv (\mathbf{M}^{*(1)})_{O_p(n^{-1/2})} + (\mathbf{M}^{*(2)})_{O_p(n^{-1})} + O_p(n^{-3/2}),$$

and $(\cdot)_{\cdot a}$ is the a-th column of a matrix with other similar notations defined similarly. Noting that

$$\text{vec}\{\mathbf{S} - \boldsymbol{\Sigma}_{\text{T}} - (\hat{\boldsymbol{\Sigma}}_{\text{NGLS*}} - \boldsymbol{\Sigma}_0)\}$$

$$= \text{vec}\{\mathbf{S} - \boldsymbol{\Sigma}_{\text{T}} - (\hat{\boldsymbol{\Sigma}}_{\text{NGLS*}} - \boldsymbol{\Sigma}_{\text{T}})\}$$

$$= \mathbf{D}_p \{\mathbf{s} - \boldsymbol{\sigma}_{\text{T}} - \boldsymbol{\Lambda}_0^{(1)} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}}) - \boldsymbol{\Lambda}_0^{(2)} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})^{<2>}\} + O_p(n^{-3/2})$$

$$= (\mathbf{D}_p - \mathbf{D}_p \boldsymbol{\Lambda}_0^{(1)}) (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}}) - \mathbf{D}_p \boldsymbol{\Lambda}_0^{(2)} (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})^{<2>} + O_p(n^{-3/2}),$$

(s1.2.2) becomes

$$\text{E}_g^{(\mathbf{s})} \{(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\text{NGLS*}})' (\hat{\boldsymbol{\Gamma}}_{\text{NT}}^{(\text{M})-1} - \boldsymbol{\Gamma}_{\text{NT}}^{(2)-1}) (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\text{NGLS*}})\}$$

$$= n^{-2} [(1/2) n^2 \text{E}_g^{(\mathbf{s})} \{(\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})' (\mathbf{D}_p - \mathbf{D}_p \boldsymbol{\Lambda}_0^{(1)})' (\mathbf{M}^{*(1)} + \mathbf{M}^{*(2)})$$

$$\times (\mathbf{D}_p - \mathbf{D}_p \boldsymbol{\Lambda}_0^{(1)}) (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})\}_{\to O(n^{-2})}$$

$$- n^2 \text{E}_g^{(\mathbf{s})} \{(\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})^{<2>}{}' (\mathbf{D}_p \boldsymbol{\Lambda}_0^{(2)})' \mathbf{M}^{*(1)} (\mathbf{D}_p - \mathbf{D}_p \boldsymbol{\Lambda}_0^{(1)}) (\mathbf{s} - \boldsymbol{\sigma}_{\text{T}})\}_{\to O(n^{-2})}]$$

$$+ O(n^{-3}).$$

(i) Under non-normality and $\boldsymbol{\sigma}_{\text{T}} = \boldsymbol{\sigma}_0$

$$
\begin{aligned}
&\mathrm{ELS}_{\mathrm{NTG}*} - \mathrm{EPLS}_{\mathrm{NTG}} \\
&= n^{-1}\{-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})\} \\
&+n^{-2}\underset{(\mathrm{A})}{[}\ 2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - 2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)}) \\
&\qquad - 6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}] \\
&+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)}+\mathbf{M}^{*(2)}) \\
&\qquad\qquad \times(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]} \\
&+O(n^{-3}).
\end{aligned}
$$

(ii) Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$

$$
\begin{aligned}
&\mathrm{ELS}_{\mathrm{NTG}*} - \mathrm{EPLS}_{\mathrm{NTG}} \\
&= n^{-1}(-2q) \\
&+n^{-2}\underset{(\mathrm{A})}{[}-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_{\mathrm{NT}}^{(3)}) - 6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)})\otimes\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\}] \\
&+(1/2)n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)}+\mathbf{M}^{*(2)}) \\
&\qquad\qquad \times(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]} \\
&+O(n^{-3}).
\end{aligned}
$$

S1.2.3 Bias correction of $\mathrm{LS}_{\mathrm{NTG}*}$

Recall that

$$
\begin{aligned}
\mathrm{LS}_{\mathrm{NTG}*} &= (\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*}) \\
&= (1/2)\mathrm{vec}'(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*})(\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1}\otimes\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1})\mathrm{vec}(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}) \\
&= (1/2)\mathrm{tr}[\{\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1}(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*})\}^2].
\end{aligned}
$$

Define

$$\mathrm{ALS}_{\mathrm{NTG}*} \equiv \mathrm{LS}_{\mathrm{NTG}*} + n^{-1}2q,$$

$$\mathrm{TLS}_{\mathrm{NTG}*} \equiv \mathrm{LS}_{\mathrm{NTG}*} + n^{-1}2\mathrm{tr}(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\mathbf{\Lambda}}^{(1)}\hat{\mathbf{\Gamma}}^{(2)})$$

$$= \mathrm{LS}_{\mathrm{NTG}*} + n^{-1}2\mathrm{tr}\{(\hat{\mathbf{\Delta}}'\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\mathbf{\Delta}})^{-1}\hat{\mathbf{\Delta}}'\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\mathbf{\Gamma}}^{(2)}\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\mathbf{\Delta}}\}$$

and

$$\mathrm{CALS}_{\mathrm{NTG}*} \equiv \mathrm{LS}_{\mathrm{NTG}*} + n^{-1}2q$$

$$+n^{-2}[2\mathrm{tr}(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\mathbf{\Lambda}}^{(2)}\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M3})}) + 6\mathrm{tr}[\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\mathbf{\Lambda}}^{(3)}\{\mathrm{vec}(\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})}) \otimes \hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})}\}]$$

$$-(1/2)n^2\widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\mathbf{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)} + \mathbf{M}^{*(2)})$$

$$\times(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\mathbf{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$+n^2\widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\mathbf{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\mathbf{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}],$$

where estimated values are given by $\hat{\mathbf{\theta}}_{\mathrm{NGLS}*}$, and $\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M3})}$ is defined using $\hat{\mathbf{\theta}}_{\mathrm{NGLS}*}$ as for $\hat{\mathbf{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})}$.

All the corrections in $\mathrm{ALS}_{\mathrm{NTG}*}$, $\mathrm{TLS}_{\mathrm{NTG}*}$ and $\mathrm{CALS}_{\mathrm{NTG}*}$ are valid only when $\mathbf{\sigma}_{\mathrm{T}} = \mathbf{\sigma}_0$.

Under normality and $\mathbf{\sigma}_{\mathrm{T}} = \mathbf{\sigma}_0$, $\mathrm{E}_f(\mathrm{ALS}_{\mathrm{NTG}*}) - \mathrm{EPLS}_{\mathrm{NTG}} = O(n^{-2})$ and $\mathrm{E}_f(\mathrm{CALS}_{\mathrm{NTG}*}) - \mathrm{EPLS}_{\mathrm{NTG}} = O(n^{-3})$.

Under non-normality and $\mathbf{\sigma}_{\mathrm{T}} = \mathbf{\sigma}_0$,
$\mathrm{E}_g(\mathrm{TLS}_{\mathrm{NTG}*}) - \mathrm{EPLS}_{\mathrm{NTG}} = O(n^{-2})$.

S1.3 $\mathrm{TLS}_{\mathrm{S}}$ by SLS when $\hat{\mathbf{W}}_{\mathbf{s}} = 2\mathbf{D}_p^+\{\mathrm{Diag}(\mathbf{S}) \otimes \mathrm{Diag}(\mathbf{S})\}\mathbf{D}_p^+{}'$ for covariance structures

S1.3.1 Definition

$$\mathrm{LS}_{\mathrm{S}} = (\mathbf{s}-\hat{\mathbf{\sigma}}_{\mathrm{SLS}})'\hat{\mathbf{W}}_{\mathrm{SLS}}^{-1}(\mathbf{s}-\hat{\mathbf{\sigma}}_{\mathrm{SLS}})$$

$$= (\mathbf{s}-\hat{\mathbf{\sigma}}_{\mathrm{SLS}})'(1/2)\mathbf{D}_p{}'\{\mathrm{Diag}^{-1}(\mathbf{S}) \otimes \mathrm{Diag}^{-1}(\mathbf{S})\}\mathbf{D}_p(\mathbf{s}-\hat{\mathbf{\sigma}}_{\mathrm{SLS}})$$

$$= (1/2)\mathrm{tr}[\{\mathrm{Diag}^{-1}(\mathbf{S})(\mathbf{S}-\hat{\mathbf{\Sigma}}_{\mathrm{SLS}})\}^2],$$

where $\mathrm{Diag}^{-1}(\mathbf{S}) = \{\mathrm{Diag}(\mathbf{S})\}^{-1}$. Note that

$$\mathbf{W}_{\rm SLS} = 2\mathbf{D}_p^+\{\mathrm{Diag}(\mathbf{\Sigma}_{\rm T})\otimes \mathrm{Diag}(\mathbf{\Sigma}_{\rm T}))\}\mathbf{D}_p^{+}{}'.$$

S1.3.2 Bias of $\mathrm{LS}_{\rm S}$

$$\begin{aligned}
&\mathrm{EPLS}_{\rm S} \equiv \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}[(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\rm SLS})' \\
&\qquad \times (1/2)\mathbf{D}_p'\{\mathrm{Diag}^{-1}(\mathbf{\Sigma}_{\rm T})\otimes \mathrm{Diag}^{-1}(\mathbf{\Sigma}_{\rm T})\}\mathbf{D}_p(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\rm SLS})], \\
&\mathrm{ELS}_{\rm S} \equiv \mathrm{E}_g^{(\mathbf{s})}(\mathrm{LS}_{\rm S}), \\
&\mathrm{ELS}_{\rm S} - \mathrm{EPLS}_{\rm S} = -2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})'\mathbf{W}_{\rm SLS}^{-1}(\hat{\boldsymbol{\sigma}}_{\rm SLS}-\boldsymbol{\sigma}_{\rm T})\} \\
&+\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\rm SLS})'(\hat{\mathbf{W}}_{\rm SLS}^{-1}-\mathbf{W}_{\rm SLS}^{-1})'(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\rm SLS})\}.
\end{aligned} \tag{s1.3.1}$$

The first term on the right-hand side of the last equation of (s1.3.1) is

$$\begin{aligned}
&-2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})'\mathbf{W}_{\rm SLS}^{-1}(\hat{\boldsymbol{\sigma}}_{\rm SLS}-\boldsymbol{\sigma}_{\rm T})\} \\
&= -2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})'\mathbf{W}_{\rm SLS}^{-1}(\hat{\boldsymbol{\sigma}}_{\rm SLS}-\boldsymbol{\sigma}_0)\} \\
&= -2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})'\}]_{\to O(n^{-1})+O(n^{-2})} \\
&\quad -2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})^{<2>}(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})'\}]_{\to O(n^{-2})} \\
&\quad -2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})^{<3>}(\mathbf{s}-\boldsymbol{\sigma}_{\rm T})'\}]_{\to O(n^{-2})} \\
&= -\{n^{-1}2\mathrm{tr}(\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_0^{(2)})\}_{O(n^{-1})} \\
&\quad + \underset{(\mathrm{A})}{[}\, n^{-2}2\mathrm{tr}(\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - n^{-2}2\mathrm{tr}(\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(2)}\mathbf{\Gamma}_0^{(3)}) \\
&\qquad - n^{-2}6\mathrm{tr}[\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(3)}\{\mathrm{vec}(\mathbf{\Gamma}_0^{(2)})\otimes\mathbf{\Gamma}_0^{(2)}\}]\,]_{(\mathrm{A})O(n^{-2})} + O(n^{-3}).
\end{aligned} \tag{s1.3.2}$$

The term of order $O(n^{-1})$ in (s1.3.2) is

$$\begin{aligned}
&-n^{-1}2\mathrm{tr}(\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_0^{(2)}) \\
&= -n^{-1}2\mathrm{tr}\{\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Delta}_0(\mathbf{\Delta}_0'\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Delta}_0+\mathbf{A}_{\rm D0})^{-1}\mathbf{\Delta}_0'\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Gamma}_0^{(2)}\} \\
&= -n^{-1}2\mathrm{tr}\{(\mathbf{\Delta}_0'\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Delta}_0+\mathbf{A}_{\rm D0})^{-1}\mathbf{\Delta}_0'\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Gamma}_0^{(2)}\mathbf{W}_{\rm SLS}^{-1}\mathbf{\Delta}_0\},
\end{aligned}$$

which is not equal to $-n^{-1}2q$ even under normality and $\boldsymbol{\sigma}_{\rm T} = \boldsymbol{\sigma}_0$ i.e., $\mathbf{A}_{\rm D0} = \mathbf{O}$ with

$$\mathbf{A}_{\mathrm{D0}} \equiv \sum_{a \ge b} (\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{W}_{\mathrm{SLS}}^{-1})_{\cdot ab} \frac{\partial^2 \sigma_{0ab}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0'}$$

$$= \sum_{a \ge b} (\boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{\mathrm{T}})_{ab} (\mathbf{W}_{\mathrm{SLS}}^{-1})_{ab,ab} \frac{\partial^2 \sigma_{0ab}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0'}$$

since $\mathbf{W}_{\mathrm{SLS}}$ is diagonal.

The second term on the right-hand side of the last equation of (s1.3.1) is

$$\begin{aligned} & \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(\hat{\mathbf{W}}_{\mathrm{SLS}}^{-1} - \mathbf{W}_{\mathrm{SLS}}^{-1})(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})\} \\ & = \mathrm{E}_g^{(\mathbf{s})}\{(1/2)\mathrm{vec}'(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{SLS}})\{\mathrm{Diag}^{-1}(\mathbf{S}) \otimes \mathrm{Diag}^{-1}(\mathbf{S}) \\ & \qquad - \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}})\}\mathrm{vec}(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{SLS}})\}. \end{aligned} \quad \text{(s1.3.3)}$$

Let $\mathbf{M}_{\mathrm{D}} = \mathrm{Diag}(\mathbf{S}) - \mathrm{Diag}(\boldsymbol{\Sigma}_{\mathrm{T}})$. Then,

$$\begin{aligned} \mathrm{Diag}^{-1}(\mathbf{S}) &= \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) - \mathrm{Diag}^{-2}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}} + \mathrm{Diag}^{-3}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^2 \\ & \quad - \mathrm{Diag}^{-4}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^3 + \mathrm{Diag}^{-5}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^4 + O_p(n^{-5/2}), \end{aligned}$$

$$\begin{aligned} & \mathrm{Diag}^{-1}(\mathbf{S}) \otimes \mathrm{Diag}^{-1}(\mathbf{S}) - \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \\ & = [-\textstyle\sum_{\mathrm{sym}}^2 \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \{\mathrm{Diag}^{-2}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}\}]_{O_p(n^{-1/2})} \\ & + [\{\mathrm{Diag}^{-2}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}\}^{<2>} + \textstyle\sum_{\mathrm{sym}}^2 \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \{\mathrm{Diag}^{-3}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^2\}]_{O_p(n^{-1})} \\ & + [-\textstyle\sum_{\mathrm{sym}}^2 \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \{\mathrm{Diag}^{-4}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^3\} \\ & \quad - \textstyle\sum_{\mathrm{sym}}^2 \{\mathrm{Diag}^{-2}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}\} \otimes \{\mathrm{Diag}^{-3}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^2\}]_{O_p(n^{-3/2})} \\ & + [\textstyle\sum_{\mathrm{sym}}^2 \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \{\mathrm{Diag}^{-5}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^4\} \\ & \quad + \{\mathrm{Diag}^{-3}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^2\}^{<2>} \\ & \quad + \textstyle\sum_{\mathrm{sym}}^2 \{\mathrm{Diag}^{-2}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}\} \otimes \{\mathrm{Diag}^{-4}(\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^3\}]_{O_p(n^{-2})} + O_p(n^{-5/2}) \\ & \equiv (\mathbf{M}_{\mathrm{D}}^{(1)})_{O_p(n^{-1/2})} + (\mathbf{M}_{\mathrm{D}}^{(2)})_{O_p(n^{-1})} + (\mathbf{M}_{\mathrm{D}}^{(3)})_{O_p(n^{-3/2})} + (\mathbf{M}_{\mathrm{D}}^{(4)})_{O_p(n^{-2})} + O_p(n^{-5/2}). \end{aligned}$$

Consequently, (s1.3.3) becomes

$$\mathrm{E}_g^{(\mathbf{s})}[(1/2)\mathrm{vec}'\{\mathbf{S}-\boldsymbol{\Sigma}_\mathrm{T}-(\hat{\boldsymbol{\Sigma}}_\mathrm{SLS}-\boldsymbol{\Sigma}_0)+\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0\}$$
$$\times(\mathbf{M}_\mathrm{D}^{(1)}+\mathbf{M}_\mathrm{D}^{(2)}+\mathbf{M}_\mathrm{D}^{(3)}+\mathbf{M}_\mathrm{D}^{(4)})$$
$$\times\mathrm{vec}\{\mathbf{S}-\boldsymbol{\Sigma}_\mathrm{T}-(\hat{\boldsymbol{\Sigma}}_\mathrm{SLS}-\boldsymbol{\Sigma}_0)+\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0]_{\to O(n^{-2})}+O(n^{-3})$$
$$(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0=O(1)).$$

As in (s1.1.5) and (s1.1.6), the term of order $O(n^{-1})$ in (s1.3.3) is

$$n^{-1}[(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}\mathrm{vec}(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)$$
$$+\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-1})}],$$

which becomes zero when $\boldsymbol{\sigma}_\mathrm{T}=\boldsymbol{\sigma}_0$.

The term of order $O(n^{-2})$ in (s1.3.3) is

$$n^{-2}\underset{(\mathrm{A})}{[}(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}_\mathrm{D}^{(2)}-\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}$$
$$+\mathbf{M}_\mathrm{D}^{(3)}+\mathbf{M}_\mathrm{D}^{(3)}\}_{\to O(n^{-2})}\mathrm{vec}(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)$$
$$+\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}_\mathrm{D}^{(2)}+\mathbf{M}_\mathrm{D}^{(3)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}+\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<3>}$$
$$+\mathbf{M}_\mathrm{D}^{(2)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}\}_{\to O(n^{-2})}$$
$$+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}_\mathrm{D}^{(1)}+\mathbf{M}_\mathrm{D}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]}.$$

When $\boldsymbol{\sigma}_\mathrm{T}=\boldsymbol{\sigma}_0$, the above becomes

$$n^{-2}\underset{(\mathrm{A})}{[}(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}_\mathrm{D}^{(1)}+\mathbf{M}_\mathrm{D}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]}.$$

$$
\begin{aligned}
&\mathrm{ELS_S}-\mathrm{EPLS_S}\\
&=n^{-1}[-2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})\\
&\qquad+(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}\mathrm{vec}(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)\\
&\qquad+\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-1})}]\qquad\qquad(\mathrm{s}1.3.4)\\
&+n^{-2}\underset{(\mathrm{A})}{[}\,2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)})-2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)})\\
&\qquad-6\mathrm{tr}[\mathbf{W}_{\mathrm{SLS}}^{-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}]\\
&+(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}_\mathrm{D}^{(2)}-\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}\\
&\qquad\qquad+\mathbf{M}_\mathrm{D}^{(3)}+\mathbf{M}_\mathrm{D}^{(4)}\}_{\to O(n^{-2})}\mathrm{vec}(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)\\
&\quad+\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}_\mathrm{D}^{(2)}+\mathbf{M}_\mathrm{D}^{(3)})\\
&\qquad\qquad\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\\
&-\mathrm{vec}'(\boldsymbol{\Sigma}_\mathrm{T}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}+\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<3>}\\
&\qquad\qquad+\mathbf{M}_\mathrm{D}^{(2)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}\}_{\to O(n^{-2})}\\
&+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}_\mathrm{D}^{(1)}+\mathbf{M}_\mathrm{D}^{(2)})\\
&\qquad\qquad\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\\
&-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\underset{(\mathrm{A})}{]}\\
&+O(n^{-3}).
\end{aligned}
$$

(i) Even under normality, when $\boldsymbol{\sigma}_\mathrm{T}-\boldsymbol{\sigma}_0=O(1)$, the term of order $O(n^{-1})$ in (s1.3.4) is not equal to $-n^{-1}2q$ though $\boldsymbol{\Gamma}_0^{(2)}=\boldsymbol{\Gamma}_\mathrm{NT}^{(2)}$.

(ii) Under non-normality and $\boldsymbol{\sigma}_\mathrm{T}=\boldsymbol{\sigma}_0$,

$$
\begin{aligned}
&\mathrm{ELS_S} - \mathrm{EPLS_S} \\
&= n^{-1}\{-2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_0^{(2)})\} \\
&+n^{-2} \underset{(\mathrm{A})}{[}\ 2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - 2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\mathbf{\Lambda}_0^{(2)}\mathbf{\Gamma}_0^{(3)}) \\
&\qquad - 6\mathrm{tr}[\mathbf{W}_{\mathrm{SLS}}^{-1}\mathbf{\Lambda}_0^{(3)}\{\mathrm{vec}(\mathbf{\Gamma}_0^{(2)}) \otimes \mathbf{\Gamma}_0^{(2)}\}] \\
&+(1/2)n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}_{\mathrm{D}}^{(1)} + \mathbf{M}_{\mathrm{D}}^{(2)}) \\
&\qquad\qquad \times (\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}_{\mathrm{D}}^{(1)}(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \underset{(\mathrm{A})}{]} \\
&+O(n^{-3}).
\end{aligned}
$$

(iii) Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, the term with $\mathbf{K}_{(4)}$ vanishes and $\mathbf{\Gamma}_0^{(j)}$ becomes $\mathbf{\Gamma}_{\mathrm{NT}}^{(j)}$ $(j = 2, 3)$. Then,

$$
\begin{aligned}
&\mathrm{ELS_S} - \mathrm{EPLS_S} \\
&= n^{-1}\{-2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_{\mathrm{NT}}^{(2)})\} \\
&+n^{-2} \underset{(\mathrm{A})}{[}\ -2\mathrm{tr}(\mathbf{W}_{\mathrm{SLS}}^{-1}\mathbf{\Lambda}_0^{(2)}\mathbf{\Gamma}_{\mathrm{NT}}^{(3)}) - 6\mathrm{tr}[\mathbf{W}_{\mathrm{SLS}}^{-1}\mathbf{\Lambda}_0^{(3)}\{\mathrm{vec}(\mathbf{\Gamma}_{\mathrm{NT}}^{(2)}) \otimes \mathbf{\Gamma}_{\mathrm{NT}}^{(2)}\}] \\
&+(1/2)n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}_{\mathrm{D}}^{(1)} + \mathbf{M}_{\mathrm{D}}^{(2)}) \\
&\qquad\qquad \times (\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}_{\mathrm{D}}^{(1)}(\mathbf{D}_p - \mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \underset{(\mathrm{A})}{]} \\
&+O(n^{-3}).
\end{aligned}
$$

S1.3.3 Bias correction of $\mathrm{LS_S}$

Recall that

$$
\begin{aligned}
\mathrm{LS_S} &= (\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(1/2)\mathbf{D}_p'\{\mathrm{Diag}^{-1}(\mathbf{S}) \otimes \mathrm{Diag}^{-1}(\mathbf{S})\}\mathbf{D}_p(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}}) \\
&= (1/2)\mathrm{tr}[\{\hat{\mathbf{\Sigma}}_{\mathrm{SLS}}^{-1}(\mathbf{S}-\hat{\mathbf{\Sigma}}_{\mathrm{SLS}})\}^2].
\end{aligned}
$$

Define

$$\mathrm{TLS_S} \equiv \mathrm{LS_S} + n^{-1}2\mathrm{tr}(\hat{\mathbf{W}}_{\mathrm{SLS}}^{-1}\hat{\mathbf{\Lambda}}^{(1)}\hat{\mathbf{\Gamma}}^{(2)})$$
$$= \mathrm{LS_S} + n^{-1}2\mathrm{tr}\{(\hat{\mathbf{\Delta}}'\hat{\mathbf{W}}_{\mathrm{SLS}}^{-1}\hat{\mathbf{\Delta}})^{-1}\hat{\mathbf{\Delta}}'\hat{\mathbf{W}}_{\mathrm{SLS}}^{-1}\hat{\mathbf{\Gamma}}^{(2)}\hat{\mathbf{W}}_{\mathrm{SLS}}^{-1}\hat{\mathbf{\Delta}}\}.$$

Note that $\mathrm{ALS_S}$ is not defined and $\hat{\mathbf{A}}_{\mathrm{D}}$ is not used in $\mathrm{TLS_S}$ since $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$ is assumed in $\mathrm{TLS_S}$.

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathrm{E}_f(\mathrm{TLS_S}) - \mathrm{EPLS_S} = O(n^{-2})$.
Under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathrm{E}_g(\mathrm{TLS_S}) - \mathrm{EPLS_S} = O(n^{-2})$.

S1.4 $\mathrm{ALS_{ADFG}}$ and $\mathrm{CALS_{ADFG}}$ when $\hat{\mathbf{W}}_{\mathbf{s}} = \hat{\mathbf{\Gamma}}^{(2)} = n\,\widehat{\mathrm{acov}_{\mathrm{ADF}}(\mathbf{s})}$ by ADF-GLS for covariance structures

Note that

$$\{n\,\widehat{\mathrm{acov}_{\mathrm{ADF}}(\mathbf{s})}\}_{ab,cd} = s_{abcd} - s_{ab}s_{cd}\ \ (p \ge a \ge b \ge 1;\ p \ge c \ge d \ge 1)$$

In this subsection, $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$ is assumed under possible non-normality.

S1.4.1 Definition

$$\mathrm{LS_{ADFG}} \equiv (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'\hat{\mathbf{\Gamma}}^{(2)-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}}),\ \mathrm{ELS_{ADFG}} \equiv \mathrm{E}_g^{(\mathbf{s})}(\mathrm{LS_{ADFG}})$$

and $\mathrm{EPLS_{ADFG}} \equiv \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'\mathbf{\Gamma}_0^{(2)-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})\}$.

S1.4.2 Bias of $\mathrm{LS_{ADFG}}$

$$\mathrm{ELS_{ADFG}} - \mathrm{EPLS_{ADFG}} = -2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{\Gamma}_0^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}} - \boldsymbol{\sigma}_{\mathrm{T}})\}$$
$$+\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'(\hat{\mathbf{\Gamma}}^{(2)-1} - \mathbf{\Gamma}_0^{(2)-1})(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})\}. \quad \text{(s1.4.1)}$$

The first term on the right-hand side of (s1.4.1) is

$$-2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{\Gamma}_0^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}} - \boldsymbol{\sigma}_{\mathrm{T}})\}$$
$$= -2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{\Gamma}_0^{(2)-1}(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}} - \boldsymbol{\sigma}_0)\}$$
$$= -2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\mathbf{\Gamma}_0^{(2)-1}\mathbf{\Lambda}_0^{(1)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\}]_{\to O(n^{-1})+O(n^{-2})}$$
$$-2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\mathbf{\Gamma}_0^{(2)-1}\mathbf{\Lambda}_0^{(2)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\}]_{\to O(n^{-2})} \quad \text{(s1.4.2)}$$
$$-2\mathrm{E}_g^{(\mathbf{s})}[\mathrm{tr}\{\mathbf{\Gamma}_0^{(2)-1}\mathbf{\Lambda}_0^{(3)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<3>}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\}]_{\to O(n^{-2})} + O(n^{-3})$$

$$= -\{n^{-1}2\mathrm{tr}(\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})\}_{O(n^{-1})}$$
$$+ \underset{(\mathrm{A})}{[}\; n^{-2}2\mathrm{tr}(\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - n^{-2}2\mathrm{tr}(\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)})$$
$$- n^{-2}6\mathrm{tr}[\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}]\;]_{(\mathrm{A})O(n^{-2})} + O(n^{-3}).$$

The term of order $O(n^{-1})$ for (s1.4.2) is

$$-n^{-1}2\mathrm{tr}(\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}) = -n^{-1}2\mathrm{tr}(\boldsymbol{\Lambda}_0^{(1)})$$
$$= -n^{-1}2\mathrm{tr}\{\boldsymbol{\Delta}_0(\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Delta}_0)^{-1}\boldsymbol{\Delta}_0'\boldsymbol{\Gamma}_0^{(2)-1}\}$$
$$= -n^{-1}2q,$$

which holds under possible non-normality. Note that there is no sample counterpart of $\mathbf{A}_0$ due to the assumption $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$.

In the second term $\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'(\hat{\boldsymbol{\Gamma}}^{(2)-1}-\boldsymbol{\Gamma}_0^{(2)-1})(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})\}$ on the right-hand side of (s1.4.1), we have

$$\hat{\boldsymbol{\Gamma}}^{(2)-1} - \boldsymbol{\Gamma}_0^{(2)-1} = [-\sum_{a\ge b}\sum_{c\ge d}(\boldsymbol{\Gamma}_0^{(2)-1})_{\cdot ab}(\boldsymbol{\Gamma}_0^{(2)-1})_{cd\cdot}\{s_{abcd} - \sigma_{\mathrm{T}abcd}$$
$$-(s_{ab}-\sigma_{\mathrm{T}ab})\sigma_{\mathrm{T}cd} - (s_{cd}-\sigma_{\mathrm{T}cd})\sigma_{\mathrm{T}ab}\}]_{O_p(n^{-1/2})}$$
$$+ \underset{(\mathrm{A})}{[}\sum_{a\ge b}\sum_{c\ge d}(\boldsymbol{\Gamma}_0^{(2)-1})_{\cdot ab}(\boldsymbol{\Gamma}_0^{(2)-1})_{cd\cdot}(s_{ab}-\sigma_{\mathrm{T}ab})(s_{cd}-\sigma_{\mathrm{T}cd})$$
$$+\sum_{a\ge b}\sum_{c\ge d}\sum_{e\ge f}\sum_{g\ge h}(\boldsymbol{\Gamma}_0^{(2)-1})_{\cdot ab}(\boldsymbol{\Gamma}_0^{(2)-1})_{cd,ef}(\boldsymbol{\Gamma}_0^{(2)-1})_{gh\cdot}$$
$$\times\{s_{abcd} - \sigma_{\mathrm{T}abcd} - (s_{ab}-\sigma_{\mathrm{T}ab})\sigma_{\mathrm{T}cd} - (s_{cd}-\sigma_{\mathrm{T}cd})\sigma_{\mathrm{T}ab}\}$$
$$\times\{s_{efgh} - \sigma_{\mathrm{T}efgh} - (s_{ef}-\sigma_{\mathrm{T}ef})\sigma_{\mathrm{T}gh} - (s_{gh}-\sigma_{\mathrm{T}gh})\sigma_{\mathrm{T}ef}\}\;]_{(\mathrm{A})O_p(n^{-1})}$$
$$+O_p(n^{-3/2})$$
$$\equiv (\mathbf{M}_{\mathrm{ADF}}^{(1)})_{O_p(n^{-1/2})} + (\mathbf{M}_{\mathrm{ADF}}^{(2)})_{O_p(n^{-1})} + O_p(n^{-3/2}).$$

Then, the second term on the right-hand side of (s1.4.1) is

$$\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'(\hat{\boldsymbol{\Gamma}}^{(2)-1}-\boldsymbol{\Gamma}_0^{(2)-1})(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})\}$$
$$=\mathrm{E}_g^{(\mathbf{s})}[\{\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}}-(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}}-\boldsymbol{\sigma}_0)\}'\{\mathbf{M}_{\mathrm{ADF}}^{(1)}+\mathbf{M}_{\mathrm{ADF}}^{(2)}+O_p(n^{-3/2})\}$$
$$\times\{\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}}-(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}}-\boldsymbol{\sigma}_0)\}]$$
$$=n^{-2}[n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{I}_{(p*)}-\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}_{\mathrm{ADF}}^{(1)}+\mathbf{M}_{\mathrm{ADF}}^{(2)})$$
$$\times(\mathbf{I}_{(p*)}-\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \qquad \text{(s1.4.3)}$$
$$-2n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'\boldsymbol{\Lambda}_0^{(2)}{}'\mathbf{M}_{\mathrm{ADF}}^{(1)}(\mathbf{I}_{(p*)}-\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}]$$
$$+O(n^{-3}),$$

where $\mathbf{M}_{\mathrm{ADF}}^{(j)}(j=1,2)$ are $p^*\times p^*$ matrices rather than $p^2\times p^2$ shown earlier.

In (s1.4.3), the following results are required:

$$n\mathrm{E}_g^{(\mathbf{s})}\{(s_{abcd}-\sigma_{\mathrm{T}abcd})(s_{ef}-\sigma_{\mathrm{T}ef})\}_{\to O(n^{-1})}$$
$$=\sigma_{\mathrm{T}abcdef}-\sigma_{\mathrm{T}abcd}\sigma_{\mathrm{T}ef}-\sum^{4}\sigma_{\mathrm{T}aef}\sigma_{\mathrm{T}bcd}$$

(Ogasawara, 2010, Subsection 1.2),

$$n\mathrm{E}_g^{(\mathbf{s})}\{(s_{abcd}-\sigma_{\mathrm{T}abcd})(s_{efgh}-\sigma_{\mathrm{T}efgh})\}_{\to O(n^{-1})}$$
$$=\sigma_{\mathrm{T}abcdefgh}-\sum^{4}(\sigma_{\mathrm{T}abcde}\sigma_{\mathrm{T}fgh}+\sigma_{\mathrm{T}efgha}\sigma_{\mathrm{T}bcd})$$
$$-\sigma_{\mathrm{T}abcd}\sigma_{\mathrm{T}efgh}+\sum^{16}\sigma_{\mathrm{T}bcd}\sigma_{\mathrm{T}fgh}\sigma_{\mathrm{T}ae}$$

(Ogasawara, 2010, Subsection 1.3)

$$n^2 \mathrm{E}_g^{(\mathbf{s})} \{(s_{abcd} - \sigma_{\mathrm{T}abcd})(s_{ef} - \sigma_{\mathrm{T}ef}))(s_{gh} - \sigma_{\mathrm{T}gh})\}_{\to O(n^{-2})}$$
$$= \sigma_{\mathrm{T}abcdefgh} - (\sigma_{\mathrm{T}abcdef}\sigma_{\mathrm{T}gh} + \sigma_{\mathrm{T}abcdgh}\sigma_{\mathrm{T}ef})$$
$$-\sum^{4} (\sigma_{\mathrm{T}bcdef}\sigma_{\mathrm{T}agh} + \sigma_{\mathrm{T}bcdgh}\sigma_{\mathrm{T}aef} + \sigma_{\mathrm{T}aefgh}\sigma_{\mathrm{T}bcd})$$
$$-\sum^{4} \sigma_{\mathrm{T}abcde}\sigma_{\mathrm{T}fgh} - 5\sigma_{\mathrm{T}abcd}\sigma_{\mathrm{T}efgh} + 6\sigma_{\mathrm{T}abcd}\sigma_{\mathrm{T}ef}\sigma_{\mathrm{T}gh}$$
$$-\sum^{4} (\sigma_{\mathrm{T}aef}\sigma_{\mathrm{T}gh} + \sigma_{\mathrm{T}agh}\sigma_{\mathrm{T}ef})\sigma_{\mathrm{T}bcd}$$
$$+\sum^{4} (\sigma_{\mathrm{T}ag}\sigma_{\mathrm{T}efh} + \sigma_{\mathrm{T}ah}\sigma_{\mathrm{T}efg} + \sigma_{\mathrm{T}ae}\sigma_{\mathrm{T}ghf} + \sigma_{\mathrm{T}af}\sigma_{\mathrm{T}ghe})\sigma_{\mathrm{T}bcd}$$
$$+\sum^{{}_4C_2=6} \{(\sigma_{\mathrm{T}aef}\sigma_{\mathrm{T}bgh} + \sigma_{\mathrm{T}agh}\sigma_{\mathrm{T}bef})\sigma_{\mathrm{T}cd} + (\sigma_{\mathrm{T}acd}\sigma_{\mathrm{T}bgh} + \sigma_{\mathrm{T}agh}\sigma_{\mathrm{T}bcd})\sigma_{\mathrm{T}ef}$$
$$+ (\sigma_{\mathrm{T}acd}\sigma_{\mathrm{T}bef} + \sigma_{\mathrm{T}aef}\sigma_{\mathrm{T}bcd})\sigma_{\mathrm{T}gh}\}$$
$$+2\sum^{3} \sigma_{\mathrm{T}ab}\sigma_{\mathrm{T}cd}(\sigma_{\mathrm{T}efgh} - \sigma_{\mathrm{T}ef}\sigma_{\mathrm{T}gh})$$

(Ogasawara, 2010, Subsection 2.1).

Then,

$$\mathrm{ELS}_{\mathrm{ADFG}} - \mathrm{EPLS}_{\mathrm{ADFG}}$$
$$= n^{-1}(-2q) + n^{-2} \underset{(\mathrm{A})}{[} 2\mathrm{tr}(\mathbf{\Gamma}_0^{(2)-1}\mathbf{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - 2\mathrm{tr}(\mathbf{\Gamma}_0^{(2)-1}\mathbf{\Lambda}_0^{(2)}\mathbf{\Gamma}_0^{(3)})$$
$$- 6\mathrm{tr}[\mathbf{\Gamma}_0^{(2)-1}\mathbf{\Lambda}_0^{(3)}\{\mathrm{vec}(\mathbf{\Gamma}_0^{(2)}) \otimes \mathbf{\Gamma}_0^{(2)}\}]$$
$$+n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})'(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)})$$
$$\times (\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \qquad \text{(s1.4.4)}$$
$$-2n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'\mathbf{\Lambda}_0^{(2)}{}'\mathbf{M}_{\mathrm{ADF}}^{(1)}(\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \underset{(\mathrm{A})}{]}$$
$$+O(n^{-3}),$$

which holds under possible non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$. Under normality, the term with $\mathbf{K}_{(4)}$ vanishes and $\mathbf{\Gamma}_0^{(j)}$ becomes $\mathbf{\Gamma}_{\rm NT}^{(j)}$ $(j = 2, 3)$.

S1.4.3 Bias correction of $\mathrm{LS}_{\rm ADFG}$

Recall that $\mathrm{LS}_{\rm ADFG} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\rm AGLS})'\hat{\mathbf{\Gamma}}^{(2)-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\rm AGLS})$. Define $\mathrm{ALS}_{\rm ADFG} \equiv \mathrm{LS}_{\rm ADFG} + n^{-1}2q$ (note that $\mathrm{TLS}_{\rm ADFG}$ is unnecessary) and

$$
\begin{aligned}
&\mathrm{CALS}_{\rm ADFG}\\
&= \mathrm{LS}_{\rm ADFG} + n^{-1}2q - n^{-2} \underset{\rm (A)}{[}\ 2\mathrm{tr}(\hat{\mathbf{\Gamma}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(1)}\hat{\mathbf{K}}_{(4)}) - 2\mathrm{tr}(\hat{\mathbf{\Gamma}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(2)}\hat{\mathbf{\Gamma}}^{(3)})\\
&\qquad - 6\mathrm{tr}[(\hat{\mathbf{\Gamma}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(3)}\{\mathrm{vec}(\hat{\mathbf{\Gamma}}^{(2)}) \otimes \hat{\mathbf{\Gamma}}^{(2)}\}]\\
&+n^2 \widehat{\mathrm{E}_g^{(\mathbf{s})}}\{(\mathbf{s} - \boldsymbol{\sigma}_{\rm T})'(\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})'(\mathbf{M}_{\rm ADF}^{(1)} + \mathbf{M}_{\rm ADF}^{(2)})\\
&\qquad\qquad \times (\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})(\mathbf{s} - \boldsymbol{\sigma}_{\rm T})\}_{\to O(n^{-2})}\\
&-2n^2 \widehat{\mathrm{E}_g^{(\mathbf{s})}}\{(\mathbf{s} - \boldsymbol{\sigma}_{\rm T})^{<2>}{}'\mathbf{\Lambda}_0^{(2)}{}'\mathbf{M}_{\rm ADF}^{(1)}(\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})(\mathbf{s} - \boldsymbol{\sigma}_{\rm T})\}_{\to O(n^{-2})} \underset{\rm (A)}{]}.
\end{aligned}
$$

Under possible non-normality and $\boldsymbol{\sigma}_{\rm T} = \boldsymbol{\sigma}_0$, we have

$$\mathrm{E}_g(\mathrm{ALS}_{\rm ADFG}) - \mathrm{EPLS}_{\rm ADFG} = O(n^{-2})$$

and $\mathrm{E}_g(\mathrm{CALS}_{\rm ADFG}) - \mathrm{EPLS}_{\rm ADFG} = O(n^{-3})$.

S1.5 $\mathrm{ALS}_{\boldsymbol{\rho}\rm ADFG}$ by ADF-GLS using $\hat{\mathbf{\Gamma}}_{\boldsymbol{\rho}}^{(2)} = n\,\widehat{\mathrm{acov}_{\rm ADF}(\mathbf{r})}$ for correlation structures

S1.5.1 Definition

Define $\mathrm{LS}_{\boldsymbol{\rho}\rm ADFG} \equiv (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\rm AGLS})'\hat{\mathbf{\Gamma}}_{\boldsymbol{\rho}}^{(2)-1}(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\rm AGLS})$, where $\mathbf{r} = \mathrm{vb}(\mathbf{R})$ is a $\{(p^2 - p)/2\} \times 1$ vector, $\mathrm{vb}(\cdot)$ is the vectorizing operator taking the off-diagonal elements below the main diagonals in a symmetric matrix, $\boldsymbol{\rho} = \mathrm{vb}(\mathbf{P})$, $\boldsymbol{\rho}_0 = \mathrm{vb}(\mathbf{P}_0)$ and $\mathbf{P}_0 = \mathbf{P}(\boldsymbol{\theta}_{\boldsymbol{\rho}0})$ is the population correlation matrix given by a structural correlation model.

$$(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}}^{(2)})_{ab,\,cd} = r_{abcd} + (1/4)r_{ab}r_{cd}(r_{aacc} + r_{bbcc} + r_{aadd} + r_{bbdd}) - (1/2)r_{ab}(r_{aacd} + r_{bbcd}) - (1/2)r_{cd}(r_{abcc} + r_{abdd}),$$

$$r_{abcd} \equiv s_{abcd} / (s_{aa}s_{bb}s_{cc}s_{dd})^{1/2} \ (p \ge a > b \ge 1;\ p \ge c > d \ge 1).$$

$\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)} = n\ \mathrm{acov}_{\mathrm{ADF}}(\mathbf{r})$ was given by Isserlis (1916, Equation (21)), Hsu, 1949, Equation (79)) and Steiger and Hakstian (1982, Equation (3.4)) (see also Steiger & Hakstian, 1983; Ogasawara, 2002, 2008).

Define $\mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{ADFG}} \equiv \mathrm{E}_g^{(\mathbf{r}^*)}\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1}(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})\}$, where $\mathbf{r}^*$ is an independent copy of $\mathbf{r}$. Let $\mathbf{P}_{\mathrm{T}}$ be the true population correlation matrix. In this subsection, $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$ is assumed with $\boldsymbol{\rho}_{\mathrm{T}} = \mathrm{vb}(\mathbf{P}_{\mathrm{T}})$.

S1.5.2 Bias of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}}$

Let $\boldsymbol{\Delta}_{\boldsymbol{\rho}0} \equiv \dfrac{\partial \boldsymbol{\rho}_0}{\partial \boldsymbol{\theta}_{\boldsymbol{\rho}}'}$, then

$$\begin{aligned}
&\mathrm{E}_g(\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{ADFG}} \\
&= -2\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r} - \boldsymbol{\rho}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1}(\hat{\boldsymbol{\rho}}_{\mathrm{AGLS}} - \boldsymbol{\rho}_{\mathrm{T}})\} \\
&\quad + \mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})'(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}}^{(2)-1} - \boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1})(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})\} \\
&= -2\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r} - \boldsymbol{\rho}_{\mathrm{T}})'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1}(\hat{\boldsymbol{\rho}}_{\mathrm{AGLS}} - \boldsymbol{\rho}_{\mathrm{T}})\}_{\to O(n^{-1})} + O(n^{-2}) \\
&= -n^{-1}2\{\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1} n\,\mathrm{acov}_{\mathrm{ADF}}(\mathbf{r})\} + O(n^{-2}) \\
&= -n^{-1}2\{(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}\} + O(n^{-2}) \\
&= -n^{-1}2q + O(n^{-2}),
\end{aligned}$$

which holds under possible non-normality.

S1.5.3 Bias correction of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}}$

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}} \equiv (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}}^{(2)-1}(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})$. Define $\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{ADFG}} = \mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}} + n^{-1}2q$ ($\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{ADFG}}$ is unnecessary while $\mathrm{CALS}_{\boldsymbol{\rho}\mathrm{ADFG}}$ can be defined but not given here). Then,

$$\mathrm{E}_g^{(\mathbf{r})}(\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{ADFG}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{ADFG}} = O(n^{-2})$$

holds under possible non-normality and $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$.

S1.6 $\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{NTG}}$ by NT-GLS using $\hat{\boldsymbol{\Gamma}}^{(2)}_{\boldsymbol{\rho}\mathrm{NT}} = n\,\widehat{\mathrm{acov}_{\mathrm{NT}}(\mathbf{r})}$ for correlation structures

S1.6.1 Definition

$\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}} \equiv (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})$, where

$$(\hat{\boldsymbol{\Gamma}}^{(2)}_{\boldsymbol{\rho}\mathrm{NT}})_{ab,\,cd} = (1/2)r_{ab}r_{cd}(r_{ac}^2 + r_{ad}^2 + r_{bc}^2 + r_{bd}^2) + r_{ac}r_{bd} + r_{ad}r_{bc} - r_{ab}(r_{bc}r_{bd} + r_{ac}r_{ad}) - r_{cd}(r_{bc}r_{ac} + r_{bd}r_{ad})$$

$(p \ge a > b \ge 1;\ p \ge c > d \ge 1)$,

$\boldsymbol{\Gamma}^{(2)}_{\boldsymbol{\rho}\mathrm{NT}} = n\ \mathrm{acov}_{\mathrm{NT}}(\mathbf{r})$ was given by Pearson and Filon (1898, Equation (xl.)), Girshick (1939, Equation (3.23)), Hsu (1949, p.400), Olkin and Siotani (1976, Equation (3.1)) and Steiger and Hakstian (1982, Equation (4.2)) (see also Ogasawara, 2002, 2008).

In this subsection, $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$ is assumed.

S1.6.2 Bias of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}$

$$\begin{aligned}
&\mathrm{E}_g(\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}} \\
&= -2\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r} - \boldsymbol{\rho}_{\mathrm{T}})'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}(\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}} - \boldsymbol{\rho}_{\mathrm{T}}) \\
&\quad + \mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'(\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}} - \hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}})(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})\} \\
&= -2\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r} - \boldsymbol{\rho}_{\mathrm{T}})'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}(\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}} - \boldsymbol{\rho}_{\mathrm{T}})\}_{\to O(n^{-1})} + O(n^{-2}) \\
&= -n^{-1}2\{\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}n\,\mathrm{acov}_{\mathrm{ADF}}(\mathbf{r})\} + O(n^{-2}) \\
&= -n^{-1}2\{(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Gamma}^{(2)}_{\boldsymbol{\rho}}\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}\} + O(n^{-2}),
\end{aligned}$$

which becomes $-n^{-1}2q$ under normality and $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$.

S1.6.3 Bias correction of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}$

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}} = (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})$. Define

$\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{NTG}} = \mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}} + n^{-1}2q$ and

$\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{NTG}} = \mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}} + n^{-1}2\mathrm{tr}\{(\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}})^{-1}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}}^{(2)}\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}\}$

($\mathrm{CALS}_{\boldsymbol{\rho}\mathrm{NTG}}$ can be defined but not given here).

Then, under normality and $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$,

$$\mathrm{E}_f^{(\mathbf{r})}(\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{NTG}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}} = O(n^{-2}),$$

$$\mathrm{E}_f^{(\mathbf{r})}(\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{NTG}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}} = O(n^{-2}).$$

Under non-normality and $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$,

$$\mathrm{E}_g^{(\mathbf{r})}(\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{NTG}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}} = O(n^{-2}).$$

S1.7 $\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{U}}$ by ULS for correlation structures

S1.7.1 Definition

$$\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}} \equiv (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{NGLS}}) = (1/2)\mathrm{tr}\{(\mathbf{R} - \hat{\mathbf{P}}_{\mathrm{ULS}})^2\}.$$

We assume that $\mathrm{Diag}(\hat{\mathbf{P}}_{\mathrm{ULS}}) = \mathbf{I}_{(p)}$. Define

$$\mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{U}} \equiv \mathrm{E}_g^{(\mathbf{r}^*)}\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{ULS}})'(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{ULS}})\}\}.$$

In this subsection, $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$ is assumed.

S1.7.2 Bias of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}}$

$$\begin{aligned}
&\mathrm{E}_g(\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{U}} \\
&= -2\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r} - \boldsymbol{\rho}_{\mathrm{T}})'(\hat{\boldsymbol{\rho}}_{\mathrm{ULS}} - \boldsymbol{\rho}_{\mathrm{T}})\} \\
&= -n^{-1}2\mathrm{tr}\{\boldsymbol{\Delta}_{\boldsymbol{\rho}0}(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}' n\,\mathrm{acov}_{\mathrm{ADF}}(\mathbf{r})\} + O(n^{-2}) \\
&= -n^{-1}2\mathrm{tr}\{(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}\} + O(n^{-2}).
\end{aligned}$$

S1.7.3 Bias correction of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}}$

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}} = (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{ULS}})'(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{ULS}})$. Define

$\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{U}} = \mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}} + n^{-1}2\mathrm{tr}\{(\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}'\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}})^{-1}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}}^{(2)}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}\}$ (note that $\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{U}}$ is not defined).

Under possible non-normality and $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$,

$$\mathrm{E}_g^{(\mathbf{r})}(\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{U}}) - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{U}} = O(n^{-2}).$$

S2. Higher-order bias corrections for cross validation criteria

S2.1 $\mathrm{ALS}_{\mathrm{NTG}}$, $\mathrm{TLS}_{\mathrm{NTG}}$ and $\mathrm{CALS}_{\text{CV-NTG}}$ when $\hat{\mathbf{W}}_{\mathbf{s}} = n\,\widehat{\mathrm{acov}_{\mathrm{NT}}(\mathbf{s})}$ by NT-GLS for covariance structures

S2.1.1 Definition

Recall that

$$\mathrm{LS}_{\mathrm{NTG}} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\hat{\mathbf{W}}_{\mathrm{NT},\mathbf{s}}^{-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}) = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})$$
$$= (1/2)\mathrm{tr}[\{\mathbf{S}^{-1}(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})\}^2] = (1/2)\mathrm{tr}\{(\mathbf{I}_{(p)} - \mathbf{S}^{-1}\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})^2\},$$

where $\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)} = \hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{s}}^{(2)} = \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}|_{\boldsymbol{\sigma}_{\mathrm{T}}=\mathbf{s}}$. Define

$$\mathrm{CV}_{\mathrm{NGLS}} = (\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})$$
$$= (1/2)\mathrm{tr}[\{\mathbf{T}^{-1}(\mathbf{T} - \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})\}^2] = (1/2)\mathrm{tr}\{(\mathbf{I}_{(p)} - \mathbf{T}^{-1}\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})^2\}$$

and

$$\mathrm{ECV}_{\mathrm{NGLS}} = \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}(\mathrm{CV}_{\mathrm{NGLS}})$$
$$= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\},$$

where the subscript **t** in $\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)}$ indicates that $\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)}$ is given by **t**, which will be omitted when obvious as in $\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)} = \hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{s}}^{(2)}$.

S2.1.2 Bias of $\mathrm{LS}_{\mathrm{NTG}}$

(i) The case of $\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0 = O(1)$

$$
\begin{aligned}
&\mathrm{ECV}_{\mathrm{NTG}} - \mathrm{EPLS}_{\mathrm{NTG}} \\
&= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\} \\
&\quad -\mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}})\} \\
&= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}[\{\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}}+\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0-(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)\}' \\
&\qquad\qquad \times(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\{\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}}+\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0-(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)\}] \\
&= [(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)'\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0) \\
&\quad +2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)]_{O(n^{-1})} \\
&\quad + \underset{(\mathrm{A})}{[}\,(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)'\mathrm{E}_g^{(\mathbf{t})}\{\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\mathrm{E}_g^{(\mathbf{t})}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}\}_{\to O(n^{-2})}(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0) \\
&\quad +2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{O_p(n^{-1})+O_p(n^{-3/2})}\}_{\to O(n^{-2})}(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0) \\
&\quad -2\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)'_{\to O(n^{-1})}\,\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0) \\
&\quad +\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&\quad -2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)_{\to O(n^{-1})} \\
&\quad +\mathrm{tr}[\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})} \\
&\qquad\quad \times\mathrm{E}_g^{(\mathbf{s})}\{(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)'\}_{\to O(n^{-1})}]\,\underset{(\mathrm{A})}{]}_{O(n^{-2})} \\
&\quad +O(n^{-3}).
\end{aligned}
\qquad \text{(s2.1.1)}
$$

(ii) The case of $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$

$$\mathrm{ECV_{NTG}} - \mathrm{EPLS_{NTG}}$$
$$= \underset{(A)}{[}\, \mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$-2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}} - \boldsymbol{\sigma}_0)_{\to O(n^{-1})}$$
$$+\mathrm{tr}[\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}$$
$$\times \mathrm{E}_g^{(\mathbf{s})}\{(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}} - \boldsymbol{\sigma}_0)(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}} - \boldsymbol{\sigma}_0)'\}_{\to O(n^{-1})}]\,\underset{(A)}{]}_{O(n^{-2})}$$
$$+O(n^{-3}).$$

(iii) Evaluation of $[\cdot]_{O(n^{-1})}$ in (s2.1.1) when $\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0 = O(1)$

The first term in $[\cdot]_{O(n^{-1})}$ is

$$(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)'\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$
$$= n^{-1}\{(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}} - \boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{T}} - \boldsymbol{\Sigma}_0)\}.$$

The second term in $[\cdot]_{O(n^{-1})}$ is

$$2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$
$$= n^{-1}[n\mathrm{E}_g^{(\mathbf{s})}\{\mathrm{vec}'(\mathbf{S} - \boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}^{(1)}\}_{\to O(n^{-1})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0),$$

where note that “2” vanishes on the right-hand side of the above equation.

(iv) Evaluation of $[\cdot]_{O(n^{-2})}$ in (s2.1.1) when $\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0 = O(1)$

The first term in $[\cdot]_{O(n^{-2})}$ is

$$(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)'\mathrm{E}_g^{(\mathbf{t})}\{\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1} - \mathrm{E}_g^{(\mathbf{t})}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}\}_{\to O(n^{-2})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$
$$= n^{-2}[(1/2)(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)'\mathbf{D}_p{}'n^2\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}^{(2)} - \mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}$$
$$+\mathbf{M}^{(3)} + \mathbf{M}^{(4)}\}_{\to O(n^{-2})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)].$$

The second term in $[\cdot]_{O(n^{-2})}$ is

$$2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{O_p(n^{-1})+O_p(n^{-3/2})}\}_{\to O(n^{-2})}(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)$$
$$=n^{-2}[2n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p'(1/2)(\mathbf{M}^{(2)}+\mathbf{M}^{(3)})\}_{\to O(n^{-2})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)]$$
$$=n^{-2}[n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p'(\mathbf{M}^{(2)}+\mathbf{M}^{(3)})\}_{\to O(n^{-2})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)].$$

The third term in $[\cdot]_{O(n^{-2})}$ is

$$-2\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)'_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)$$
$$=-n^{-2}[2\{\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\}'(1/2)\mathbf{D}_p'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)]$$
$$=-n^{-2}[\{\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\}'\mathbf{D}_p'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0)]$$

The fourth term in $[\cdot]_{O(n^{-2})}$ is

$$\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$=n^{-2}[n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})(1/2)\mathbf{D}_p'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}].$$

The fifth term in $[\cdot]_{O(n^{-2})}$ is

$$-2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})\}_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)_{\to O(n^{-1})}$$
$$=-n^{-2}[2n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})(1/2)\mathbf{D}_p'\mathbf{M}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})]$$
$$=-n^{-2}[n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\mathbf{D}_p'\mathbf{M}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})].$$

The sixth term in $[\cdot]_{O(n^{-2})}$ is

$$\mathrm{tr}[\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{t}}^{(2)-1}-\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{s})}\{(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}}-\boldsymbol{\sigma}_0)'\}_{\to O(n^{-1})}]$$
$$=n^{-2}\mathrm{tr}\{(1/2)\mathbf{D}_p'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\}.$$

(v) Evaluation when $\boldsymbol{\sigma}_{\mathrm{T}}=\boldsymbol{\sigma}_0$

In $\mathrm{ECV}_{\mathrm{NTG}}-\mathrm{EPLS}_{\mathrm{NTG}}$, the term of order $O(n^{-1})$ vanishes. The term of order $O(n^{-2})$ becomes the sum of the 4th, 5th and 6th terms in (iv), which under non-normality is

$$n^{-2}\underset{(A)}{[}\ n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})(1/2)\mathbf{D}_p{}'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\mathbf{D}_p{}'\mathbf{M}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})$$

$$+\mathrm{tr}\{(1/2)\mathbf{D}_p{}'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\}\underset{(A)}{]}$$

and under normality is

$$n^{-2}\underset{(A)}{[}\ n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})(1/2)\mathbf{D}_p{}'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-n\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\mathbf{D}_p{}'\mathbf{M}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)})$$

$$+\mathrm{tr}\{(1/2)\mathbf{D}_p{}'n\mathrm{E}_f^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\}\underset{(A)}{]}.$$

Define $\mathrm{ELS}_{\mathrm{NTG}}=\mathrm{E}_g(\mathrm{LS}_{\mathrm{NTG}})$, then, we have

$$\mathrm{ELS}_{\mathrm{NTG}}-\mathrm{ECV}_{\mathrm{NGLS}}$$

$$=(\mathrm{ELS}_{\mathrm{NTG}}-\mathrm{EPLS}_{\mathrm{NTG}})_{\to O(n^{-2})}-(\mathrm{ECV}_{\mathrm{NGLS}}-\mathrm{EPLS}_{\mathrm{NTG}})_{\to O(n^{-2})}$$

$$+O(n^{-3})$$

(the first term on the right-hand side of the above equation is given by Subsection S1.1.2 and the second term is given by the negative of the preceding results in this subsection)

$$=n^{-1}[-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})$$

$$-\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-1})}] \qquad \text{(s2.1.2)}$$

(note that three terms have been canceled)

$$+n^{-2}\underset{(A)}{[}\ 2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)})-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)})$$

$$-6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}]$$

$$+\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}^{(2)}+\mathbf{M}^{(3)})$$

$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}+\mathbf{M}^{(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<3>}$$

$$+\mathbf{M}^{(2)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}\}$$

$$-n^2 \mathrm{E}_g^{(\mathbf{s})}\{\mathrm{vec}'(\mathbf{S}-\boldsymbol{\Sigma}_{\mathrm{T}})(\mathbf{M}^{(2)}+\mathbf{M}^{(3)})\}_{\to O(n^{-2})}\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0) \text{ ###}$$

$$+\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}}-\boldsymbol{\Sigma}_0)n^2 \mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)}) \text{ ###}$$

(two term have been canceled, ### indicates added terms for cross validation criteria over LS criteria)

$$+(1/2)n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$

$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-(1/2)n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p{}'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \text{ ###}$$

$$+n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p{}'\mathbf{M}^{(1)}\mathbf{D}_p\}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)}) \text{ ###}$$

$$-\mathrm{tr}\{(1/2)\}\mathbf{D}_p{}'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\} \text{ ###} \underset{(\mathrm{A})}{]}+O_p(n^{-3}).$$

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}}-\boldsymbol{\sigma}_0=O(1)$, the first term in $n^{-1}[\cdot]$ of (s2.1.2) is $-2\mathrm{tr}(\boldsymbol{\Lambda}_0^{(1)})=-2\mathrm{tr}\{(\boldsymbol{\Delta}_0{}'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0+\mathbf{A}_0)^{-1}\boldsymbol{\Delta}_0{}'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Delta}_0\}\neq -2q$, and the first term in $n^{-2}\underset{(\mathrm{A})}{[}\cdot\underset{(\mathrm{A})}{]}$ becomes $2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)})=0$.

Under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}}=\boldsymbol{\sigma}_0$,

$$\mathrm{ELS}_{\mathrm{NTG}}-\mathrm{ECV}_{\mathrm{NGLS}}$$

$$=n^{-1}\{-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})\}$$

$$+n^{-2}\underset{(\mathrm{A})}{[}\,2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)})-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)})$$

$$-6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}]$$

$$+(1/2)n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$

$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-n^2 \mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\#$$

$$+n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p'\mathbf{M}^{(1)}\mathbf{D}_p\}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\#\#\#$$

$$-\mathrm{tr}\{(1/2)\}\mathbf{D}_p'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\}\#\#\#\ \underset{(\mathrm{A})}{]}$$

$$+O_p(n^{-3}).$$

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}}=\boldsymbol{\sigma}_0$,

$$\mathrm{ELS}_{\mathrm{NTG}}-\mathrm{ECV}_{\mathrm{NGLS}}$$

$$=n^{-1}(-2q)$$

$$+n^{-2}\underset{(\mathrm{A})}{[}-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_{\mathrm{NT}}^{(3)})-6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)})\otimes\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\}]$$

$$+(1/2)n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$

$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$-(1/2)n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\#$$

$$+n\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p'\mathbf{M}^{(1)}\mathbf{D}_p\}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)})\#\#\#$$

$$-\mathrm{tr}\{(1/2)\mathbf{D}_p'n\mathrm{E}_f^{(\mathbf{s})}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\}\#\#\#\ \underset{(\mathrm{A})}{]}$$

$$+O_p(n^{-3}).$$

S2.1.3 Bias correction of $\mathrm{LS}_{\mathrm{NTG}}$

Recall that

$$\mathrm{LS}_{\mathrm{NTG}}=(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{NGLS}})=(1/2)\mathrm{tr}\{(\mathbf{I}_{(p)}-\mathbf{S}^{-1}\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}})^2\},$$

$$\mathrm{ALS}_{\mathrm{NTG}}=\mathrm{LS}_{\mathrm{NTG}}+n^{-1}2q$$

and

$$\mathrm{TLS}_{\mathrm{NTG}}=\mathrm{LS}_{\mathrm{NTG}}+n^{-1}2\mathrm{tr}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Lambda}}^{(1)}\hat{\boldsymbol{\Gamma}}^{(2)})$$

$$=\mathrm{LS}_{\mathrm{NTG}}+n^{-1}2\mathrm{tr}\{(\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Gamma}}^{(2)}\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Delta}}\}.$$

Define

$$\mathrm{CALS}_{\text{CV-NTG}} = \mathrm{LS}_{\text{NTG}} + n^{-1}2q$$
$$-n^{-2}\underset{(\mathrm{A})}{[}-2\mathrm{tr}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Lambda}}^{(2)}\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(3)}) - 6\mathrm{tr}[\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)-1}\hat{\boldsymbol{\Lambda}}^{(3)}\{\mathrm{vec}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)})\otimes\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)}\}]$$
$$+(1/2)n^2\widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$-n^2\widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$
$$-(1/2)n^2\widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p{}'(\mathbf{M}^{(1)}+\mathbf{M}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\ \#\#\#$$
$$+n\widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p{}'\mathbf{M}^{(1)}\mathbf{D}_p\}\hat{\boldsymbol{\Lambda}}^{(2)}\mathrm{vec}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)})\ \#\#\#$$
$$-\mathrm{tr}\{(1/2)\mathbf{D}_p{}'n\widehat{\mathrm{E}_f^{(\mathbf{s})}}(\mathbf{M}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\hat{\boldsymbol{\Lambda}}^{(1)}\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)}\hat{\boldsymbol{\Lambda}}^{(1)}{}'\}\ \#\#\#\ \underset{(\mathrm{A})}{]}$$
$$+O_p(n^{-3}).$$

Note that the bias corrections in $\mathrm{ALS}_{\text{NTG}}$, $\mathrm{TLS}_{\text{NTG}}$ and $\mathrm{CALS}_{\text{CV-NTG}}$ are valid only when a structural model is true i.e., $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$.

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathrm{E}_f(\mathrm{ALS}_{\text{NTG}}) - \mathrm{ECV}_{\text{NTG}} = O(n^{-2})$ and $\mathrm{E}_f(\mathrm{CALS}_{\text{NTG}}) - \mathrm{ECV}_{\text{NTG}} = O(n^{-3})$.

Under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathrm{E}_g(\mathrm{TLS}_{\text{NTG}}) - \mathrm{ECV}_{\text{NTG}} = O(n^{-2})$.

S2.2 $\mathrm{ALS}_{\text{NTG*}}$, $\mathrm{TLS}_{\text{NTG*}}$, and $\mathrm{CALS}_{\text{CV-NTG*}}$ by NT-GLS* when $\hat{\mathbf{W}}_{\mathbf{s}} = \hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})}$ ($(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})})_{ab,\,cd} = \hat{\sigma}_{\text{NTGLS*},\,ac}\hat{\sigma}_{\text{NTGLS*},\,bd} + \hat{\sigma}_{\text{NTGLS*},\,ad}\hat{\sigma}_{\text{NTGLS*},\,bc};\ p \ge a \ge b \ge 1;\ p \ge c \ge d \ge 1$) for covariance structures

S2.2.1 Definition

Recall that

$$\mathrm{LS}_{\mathrm{NTG}*} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{s}}^{(\mathrm{M})-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})$$

$$= (1/2)\mathrm{vec}'(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*})(\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1} \otimes \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1})\mathrm{vec}(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*})$$

$$= (1/2)\mathrm{tr}[\{\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1}(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*})\}^2]$$

$$= (1/2)\mathrm{tr}\{(\mathbf{I}_{(p)} - \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1}\mathbf{S})^2\}.$$

Define

$$\mathrm{ECV}_{\mathrm{NGLS}*} \equiv \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT},\mathbf{s}}^{(\mathrm{M})-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})\}$$

$$= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}[(1/2)\mathrm{tr}\{(\mathbf{I}_{(p)} - \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1}\mathbf{T})^2\}].$$

In this subsection, $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$ is assumed unless otherwise stated.

S2.2.2 Bias of $\mathrm{LS}_{\mathrm{NTG}*}$

Recall that $\mathrm{ELS}_{\mathrm{NTG}*} = \mathrm{E}_g(\mathrm{LS}_{\mathrm{NTG}*})$. Then,

$$\mathrm{ELS}_{\mathrm{NTG}*} - \mathrm{ECV}_{\mathrm{NGLS}*}$$

$$= (\mathrm{ELS}_{\mathrm{NTG}*} - \mathrm{EPLS}_{\mathrm{NTG}*}) - (\mathrm{ECV}_{\mathrm{NGLS}*} - \mathrm{EPLS}_{\mathrm{NTG}*}),$$

where the first term on the right-hand side of the above equation was given in Subsection S1.2.2.

Under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$,

$$\mathrm{ECV}_{\mathrm{NGLS}*} - \mathrm{EPLS}_{\mathrm{NTG}}$$

$$= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT,\,s}}^{(\mathrm{M})-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})\}$$

$$- \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})'\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})\}$$

$$= \underset{(\mathrm{A})}{[}\ \mathrm{tr}[\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT,\,s}}^{(\mathrm{M})-1} - \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}})(\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}})'\}_{\to O(n^{-1})}]$$

$$+\mathrm{E}_g^{(\mathbf{s})}\{(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*} - \boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT,\,s}}^{(\mathrm{M})-1} - \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1})(\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\ \underset{(\mathrm{A})}{]}_{O(n^{-2})}$$

$$+ O(n^{-3})$$

$$= n^{-2}\underset{(\mathrm{A})}{[}\ (1/2)\mathrm{tr}\{\mathbf{D}_p{}'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{*(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Gamma}_0^{(2)}\}$$

$$+ (1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)} + \mathbf{M}^{*(2)})\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$+n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\ \underset{(\mathrm{A})}{]} + O(n^{-3}).$$

(s2.2.1)

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$,

$$
\begin{aligned}
&\mathrm{ECV}_{\mathrm{NGLS}*} - \mathrm{EPLS}_{\mathrm{NTG}} \\
&= n^{-2} \underset{(\mathrm{A})}{[} (1/2)\mathrm{tr}\{\mathbf{D}_p{}' n\mathrm{E}_f^{(\mathbf{s})}(\mathbf{M}^{*(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\} \\
&\quad + (1/2)n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)}+\mathbf{M}^{*(2)})\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&+ n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \underset{(\mathrm{A})}{]} + O(n^{-3}).
\end{aligned}
$$

(s2.2.2)

Then, under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, from (i) of Subsection S1.2.2 and (s2.2.1),

$$
\begin{aligned}
&\mathrm{ELS}_{\mathrm{NTG}*} - \mathrm{ECV}_{\mathrm{NGLS}*} \\
&= n^{-1}\{-2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})\} \\
&+ n^{-2} \underset{(\mathrm{A})}{[} 2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - 2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)}) \\
&\quad - 6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}] \\
&+ (1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)}+\mathbf{M}^{*(2)}) \\
&\qquad\qquad \times(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&- n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}(\mathbf{D}_p-\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&- (1/2)\mathrm{tr}\{\mathbf{D}_p{}' n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}^{*(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Gamma}_0^{(2)}\}\#\#\# \\
&- (1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)}+\mathbf{M}^{*(2)}) \\
&\qquad\qquad \times\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\# \\
&- n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)} \\
&\qquad\qquad \times\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\# \underset{(\mathrm{A})}{]} + O(n^{-3}).
\end{aligned}
$$

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, from (ii) of Subsection S1.2.2 and (s2.2.2),

$$
\begin{aligned}
&\mathrm{ELS}_{\mathrm{NTG}*} - \mathrm{ECV}_{\mathrm{NGLS}*} \\
&= n^{-1}(-2q) \\
&+ n^{-2} \underset{(\mathrm{A})}{[} -2\mathrm{tr}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_{\mathrm{NT}}^{(3)}) - 6\mathrm{tr}[\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}) \otimes \boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\}] \\
&+ (1/2)n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)} + \mathbf{M}^{*(2)}) \\
&\qquad\qquad \times (\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&- n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&- (1/2)\mathrm{tr}\{\mathbf{D}_p{}' n\mathrm{E}_f^{(\mathbf{s})}(\mathbf{M}^{*(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Gamma}_{\mathrm{NT}}^{(2)}\}\#\#\# \\
&- (1/2)n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)} + \mathbf{M}^{*(2)}) \\
&\qquad\qquad\qquad \times \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\# \\
&- n^2 \mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)} \\
&\qquad\qquad\qquad \times \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\# \underset{(\mathrm{A})}{]} + O(n^{-3}).
\end{aligned}
$$

S2.2.3 Bias correction of $\mathrm{LS}_{\mathrm{NTG}*}$

Recall that

$$
\begin{aligned}
&\mathrm{LS}_{\mathrm{NTG}*} = (\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*})'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}(\mathbf{s}-\hat{\boldsymbol{\sigma}}_{\mathrm{NGLS}*}) \\
&= (1/2)\mathrm{vec}'(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*})(\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1} \otimes \hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1})\mathrm{vec}(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}) \\
&= (1/2)\mathrm{tr}[\{\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*}^{-1}(\mathbf{S}-\hat{\boldsymbol{\Sigma}}_{\mathrm{NGLS}*})\}^2],
\end{aligned}
$$

$\mathrm{ALS}_{\mathrm{NTG}*} = \mathrm{LS}_{\mathrm{NTG}*} + n^{-1}2q$ and

$$
\begin{aligned}
\mathrm{TLS}_{\mathrm{NTG}*} &= \mathrm{LS}_{\mathrm{NTG}*} + n^{-1}2\mathrm{tr}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\boldsymbol{\Lambda}}^{(1)}\hat{\boldsymbol{\Gamma}}^{(2)}) \\
&= \mathrm{LS}_{\mathrm{NTG}*} + n^{-1}2\mathrm{tr}\{(\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\boldsymbol{\Gamma}}^{(2)}\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\boldsymbol{\Delta}}\}.
\end{aligned}
$$

Define

$$
\begin{aligned}
&\mathrm{CALS}_{\text{CV-NTG*}} \\
&= \mathrm{LS}_{\text{NTG*}} + n^{-1}(-2q) \\
&+n^{-2} \underset{(\mathrm{A})}{[} \; 2\mathrm{tr}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\boldsymbol{\Lambda}}^{(2)}\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M3})}) + 6\mathrm{tr}[\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})-1}\hat{\boldsymbol{\Lambda}}^{(3)}\{\mathrm{vec}(\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})}) \otimes \hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(\mathrm{M})}\}] \\
&-(1/2)n^2 \widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)} + \mathbf{M}^{*(2)}) \\
&\qquad\qquad \times (\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&+n^2 \widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)}(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&+(1/2)\mathrm{tr}\{\mathbf{D}_p{}' n\widehat{\mathrm{E}_f^{(\mathbf{s})}}(\mathbf{M}^{*(2)})_{\to O(n^{-1})}\mathbf{D}_p\hat{\boldsymbol{\Gamma}}_{\mathrm{NT}}^{(2)}\}\#\#\# \\
&+(1/2)n^2 \widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}^{*(1)} + \mathbf{M}^{*(2)}) \\
&\qquad\qquad\qquad \times \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\# \\
&+n^2 \widehat{\mathrm{E}_f^{(\mathbf{s})}}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}^{*(1)} \\
&\qquad\qquad\qquad \times \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}\#\#\# \underset{(\mathrm{A})}{]} + O(n^{-3}).
\end{aligned}
$$

All the corrections in $\mathrm{ALS}_{\text{NTG*}}$, $\mathrm{TLS}_{\text{NTG*}}$ and $\mathrm{CALS}_{\text{CV-NTG*}}$ are valid only when $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$.

Under normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$, $\mathrm{E}_f(\mathrm{ALS}_{\text{NTG*}}) - \mathrm{ECV}_{\text{NGLS*}} = O(n^{-2})$ and $\mathrm{E}_f(\mathrm{CALS}_{\text{CV-NTG*}}) - \mathrm{ECV}_{\text{NGLS*}} = O(n^{-3})$.

Under non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$,
$\mathrm{E}_g(\mathrm{TLS}_{\text{NTG*}}) - \mathrm{ECV}_{\text{NGLS*}} = O(n^{-2})$.

S2.3 $\mathrm{TLS}_{\mathrm{S}}$ by SLS when $\hat{\mathbf{W}}_{\mathbf{s}} = 2\mathbf{D}_p^{+}\{\mathrm{Diag}(\mathbf{S}) \otimes \mathrm{Diag}(\mathbf{S})\}\mathbf{D}_p^{+}{}'$ for covariance structures

S2.3.1 Definition

Recall that

$$\mathrm{LS_S} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})' \hat{\mathbf{W}}_{\mathrm{SLS}}^{-1} (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})$$
$$= (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(1/2)\mathbf{D}_p '\{\mathrm{Diag}^{-1}(\mathbf{S}) \otimes \mathrm{Diag}^{-1}(\mathbf{S})\}\mathbf{D}_p (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})$$
$$= (1/2)\mathrm{tr}[\{\mathrm{Diag}^{-1}(\mathbf{S})(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{SLS}})\}^2]$$
and $\mathrm{ELS_S} = \mathrm{E}_g(\mathrm{LS_S})$.

S2.3.2 Bias of $\mathrm{LS_S}$

Define

$$\mathrm{ECV_{SLS}} = \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}[(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(1/2)\mathbf{D}_p ' \times \{\mathrm{Diag}^{-1}(\mathbf{T}) \otimes \mathrm{Diag}^{-1}(\mathbf{T})\}\mathbf{D}_p (\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})].$$

Then,

$\mathrm{ELS_S} - \mathrm{ECV_{SLS}} = (\mathrm{ELS_S} - \mathrm{EPLS_S}) - (\mathrm{ECV_{SLS}} - \mathrm{EPLS_S})$, where the first term was given by Subsection 1.3.2. Let

$\hat{\mathbf{V}}_{\mathbf{t}}^{-1} \equiv (1/2)\mathbf{D}_p '\{\mathrm{Diag}^{-1}(\mathbf{T}) \otimes \mathrm{Diag}^{-1}(\mathbf{T})\}\mathbf{D}_p$ and

$\mathbf{V}^{-1} \equiv (1/2)\mathbf{D}_p '\{\mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}})\}\mathbf{D}_p$. Then, the negative second term is

$$\mathrm{ECV_{SLS}} - \mathrm{EPLS_S}$$
$$= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}[(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(1/2)\mathbf{D}_p ' \times \{\mathrm{Diag}^{-1}(\mathbf{T}) \otimes \mathrm{Diag}^{-1}(\mathbf{T})\}\mathbf{D}_p (\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})]$$
$$- \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}[(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(1/2)\mathbf{D}_p ' \times \{\mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}}) \otimes \mathrm{Diag}^{-1}(\boldsymbol{\Sigma}_{\mathrm{T}})\}\mathbf{D}_p (\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})]$$
$$= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})\}$$
$$= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}[\{\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}} + \boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0 - (\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}} - \boldsymbol{\sigma}_0)\}' \times (\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})\{\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}} + \boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0 - (\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}} - \boldsymbol{\sigma}_0)\}$$
$$= [(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)' \mathrm{E}_g^{(\mathbf{t})}(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$
$$+ 2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})\}_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)]_{O(n^{-1})}$$

$$+ \underset{(\mathrm{A})}{[}\ (\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)'\mathrm{E}_g^{(\mathbf{t})}\{\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathrm{E}_g^{(\mathbf{t})}(\hat{\mathbf{V}}_{\mathbf{t}}^{-1})_{\to O(n^{-1})}\}_{\to O(n^{-2})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$

$$+ 2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})_{O_p(n^{-1}) + O_p(n^{-3/2})}\}_{\to O(n^{-2})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$

$$- 2\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}} - \boldsymbol{\sigma}_0)_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{t})}(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})_{\to O(n^{-1})}(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$

$$+ \mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})(\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$- 2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t} - \boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})\}_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}} - \boldsymbol{\sigma}_0)_{\to O(n^{-1})}$$

$$+ \mathrm{tr}[\mathrm{E}_g^{(\mathbf{t})}(\hat{\mathbf{V}}_{\mathbf{t}}^{-1} - \mathbf{V}^{-1})_{\to O(n^{-1})}$$

$$\times \mathrm{E}_g^{(\mathbf{s})}\{(\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}} - \boldsymbol{\sigma}_0)(\hat{\boldsymbol{\sigma}}_{\mathrm{SLS}} - \boldsymbol{\sigma}_0)'\}_{\to O(n^{-1})}\ \underset{(\mathrm{A})}{]}_{O(n^{-2})}$$

$$+ O(n^{-3})$$

$$= n^{-1}[(1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}} - \boldsymbol{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_{\mathrm{D}}^{(2)})_{\to O(n^{-1})}\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}} - \boldsymbol{\Sigma}_0)$$

$$+ n\mathrm{E}_g^{(\mathbf{s})}\{\mathrm{vec}'(\mathbf{S} - \boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{M}_{\mathrm{D}}^{(1)}\}_{\to O(n^{-1})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)]$$

$$+ n^{-2}\ \underset{(\mathrm{A})}{[}\ (1/2)\mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}} - \boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}_{\mathrm{D}}^{(2)} - \mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_{\mathrm{D}}^{(2)})_{\to O(n^{-1})}$$

$$+ \mathbf{M}_{\mathrm{D}}^{(3)} + \mathbf{M}_{\mathrm{D}}^{(4)}\}_{\to O(n^{-2})}\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{T}} - \boldsymbol{\Sigma}_0)$$

$$+ \mathrm{vec}'(\boldsymbol{\Sigma}_{\mathrm{T}} - \boldsymbol{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}_{\mathrm{D}}^{(2)} + \mathbf{M}_{\mathrm{D}}^{(3)})\mathbf{D}_p(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$- \{\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\}'\mathbf{D}_p{}'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_{\mathrm{D}}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p(\boldsymbol{\sigma}_{\mathrm{T}} - \boldsymbol{\sigma}_0)$$

$$+ n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'(1/2)\mathbf{D}_p{}'(\mathbf{M}_{\mathrm{D}}^{(1)} + \mathbf{M}_{\mathrm{D}}^{(2)})\mathbf{D}_p(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})}$$

$$- n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{D}_p{}'\mathbf{M}_{\mathrm{D}}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})$$

$$+ \mathrm{tr}\{(1/2)\mathbf{D}_p{}'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_{\mathrm{D}}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\ \underset{(\mathrm{A})}{]} + O(n^{-3}).$$

From these results,

$$\mathrm{ELS_S}-\mathrm{ECV_{SLS}}$$
$$=n^{-1}[-2\mathrm{tr}(\mathbf{V}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_0^{(2)})$$
$$-\mathrm{vec}'(\mathbf{\Sigma}_\mathrm{T}-\mathbf{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}\{\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\mathbf{\Lambda}_0^{(1)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-1})}]$$

(four terms have been canceled)

$$+n^{-2}\underset{(\mathrm{A})}{[}\;2\mathrm{tr}(\mathbf{V}^{-1}\mathbf{\Lambda}_0^{(1)}\mathbf{K}_{(4)})-2\mathrm{tr}(\mathbf{V}^{-1}\mathbf{\Lambda}_0^{(2)}\mathbf{\Gamma}_0^{(3)})$$
$$-6\mathrm{tr}[\mathbf{V}^{-1}\mathbf{\Lambda}_0^{(3)}\{\mathrm{vec}(\mathbf{\Gamma}_0^{(2)})\otimes\mathbf{\Gamma}_0^{(2)}\}]$$
$$+\mathrm{vec}'(\mathbf{\Sigma}_\mathrm{T}-\mathbf{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}_\mathrm{D}^{(2)}+\mathbf{M}_\mathrm{D}^{(3)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-\mathrm{vec}'(\mathbf{\Sigma}_\mathrm{T}-\mathbf{\Sigma}_0)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\mathbf{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}+\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\mathbf{\Lambda}_0^{(3)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<3>}$$
$$+\mathbf{M}_\mathrm{D}^{(2)}\mathbf{D}_p\mathbf{\Lambda}_0^{(2)}(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{\mathrm{vec}'(\mathbf{S}-\mathbf{\Sigma}_\mathrm{T})(\mathbf{M}_\mathrm{D}^{(2)}+\mathbf{M}_\mathrm{D}^{(3)})\}_{\to O(n^{-2})}\mathrm{vec}(\mathbf{\Sigma}_\mathrm{T}-\mathbf{\Sigma}_0)\#\#\#$$
$$+\mathrm{vec}'(\mathbf{\Sigma}_\mathrm{T}-\mathbf{\Sigma}_0)n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\mathrm{vec}(\mathbf{\Lambda}_0^{(2)})\#\#\#$$

$$+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})'(\mathbf{M}_\mathrm{D}^{(1)}+\mathbf{M}_\mathrm{D}^{(2)})$$
$$\times(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}{}'(\mathbf{D}_p\mathbf{\Lambda}_0^{(2)})'\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p-\mathbf{D}_p\mathbf{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'\mathbf{D}_p{}'(\mathbf{M}_\mathrm{D}^{(1)}+\mathbf{M}_\mathrm{D}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\#\#\#$$
$$+n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'\mathbf{D}_p{}'\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\mathbf{\Lambda}_0^{(2)}\mathrm{vec}(\mathbf{\Gamma}_0^{(2)})\#\#\#$$
$$-\mathrm{tr}\{(1/2)\mathbf{D}_p{}'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_0^{(2)}\mathbf{\Lambda}_0^{(1)}{}'\#\#\#\underset{(\mathrm{A})}{]}+O(n^{-3}).$$

When $\boldsymbol{\sigma}_\mathrm{T}-\boldsymbol{\sigma}_0=O(1)$, even under normality the first term in $n^{-1}[\cdot]$ does not become $-2q$ though $\mathbf{\Gamma}_0^{(2)}=\mathbf{\Gamma}_\mathrm{NT}^{(2)}$.

Under non-normality and $\boldsymbol{\sigma}_\mathrm{T}=\boldsymbol{\sigma}_0$,

$$\mathrm{ELS_S} - \mathrm{ECV_{SLS}}$$
$$= n^{-1}\{-2\mathrm{tr}(\mathbf{V}^{-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)})\}$$
$$+n^{-2}\underset{(\mathrm{A})}{[}\ 2\mathrm{tr}(\mathbf{V}^{-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - 2\mathrm{tr}(\mathbf{V}^{-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)})$$
$$-6\mathrm{tr}[\mathbf{V}^{-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\otimes\boldsymbol{\Gamma}_0^{(2)}\}]$$
$$+(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}_\mathrm{D}^{(1)} + \mathbf{M}_\mathrm{D}^{(2)})$$
$$\times(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-(1/2)n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'\mathbf{D}_p{}'(\mathbf{M}_\mathrm{D}^{(1)} + \mathbf{M}_\mathrm{D}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\ \#\#\#$$
$$+n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'\mathbf{D}_p{}'\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\ \#\#\#$$
$$-\mathrm{tr}\{(1/2)\mathbf{D}_p{}'n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\ \#\#\#\ \underset{(\mathrm{A})}{]} + O(n^{-3}).$$

Under normality and $\boldsymbol{\sigma}_\mathrm{T} = \boldsymbol{\sigma}_0$ with $\mathbf{K}_{(4)} = \mathbf{O}$ and $\boldsymbol{\Gamma}_0^{(2)} = \boldsymbol{\Gamma}_\mathrm{NT}^{(2)}$,

$$\mathrm{ELS_S} - \mathrm{ECV_{SLS}}$$
$$= n^{-1}\{-2\mathrm{tr}(\mathbf{V}^{-1}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_\mathrm{NT}^{(2)})\}$$
$$+n^{-2}\underset{(\mathrm{A})}{[}\ -2\mathrm{tr}(\mathbf{V}^{-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_\mathrm{NT}^{(3)}) - 6\mathrm{tr}[\mathbf{V}^{-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_\mathrm{NT}^{(2)})\otimes\boldsymbol{\Gamma}_\mathrm{NT}^{(2)}\}]$$
$$+(1/2)n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}_\mathrm{D}^{(1)} + \mathbf{M}_\mathrm{D}^{(2)})$$
$$\times(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})^{<2>}{}'(\mathbf{D}_p\boldsymbol{\Lambda}_0^{(2)})'\mathbf{M}_\mathrm{D}^{(1)}(\mathbf{D}_p - \mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}$$
$$-(1/2)n^2\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'\mathbf{D}_p{}'(\mathbf{M}_\mathrm{D}^{(1)} + \mathbf{M}_\mathrm{D}^{(2)})\mathbf{D}_p(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})\}_{\to O(n^{-2})}\ \#\#\#$$
$$+n\mathrm{E}_f^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_\mathrm{T})'\mathbf{D}_p{}'\mathbf{M}_\mathrm{D}^{(1)}\mathbf{D}_p\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)})\ \#\#\#$$
$$-\mathrm{tr}\{(1/2)\mathbf{D}_p{}'n\mathrm{E}_f^{(\mathbf{s})}(\mathbf{M}_\mathrm{D}^{(2)})_{\to O(n^{-1})}\mathbf{D}_p\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'\ \#\#\#\ \underset{(\mathrm{A})}{]} + O(n^{-3}).$$

S2.3.3 Bias correction of $\mathrm{LS_S}$

Recall that

$$\mathrm{LS_S} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'\hat{\mathbf{W}}_{\mathrm{SLS}}^{-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})$$
$$= (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'\hat{\mathbf{V}}_{\mathbf{s}}^{-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})$$
$$= (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})'(1/2)\mathbf{D}_p'\{\mathrm{Diag}^{-1}(\mathbf{S}) \otimes \mathrm{Diag}^{-1}(\mathbf{S})\}\mathbf{D}_p(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{SLS}})$$
$$= (1/2)\mathrm{tr}[\{\mathrm{Diag}^{-1}(\mathbf{S})(\mathbf{S} - \hat{\boldsymbol{\Sigma}}_{\mathrm{SLS}})\}^2].$$

Define

$$\mathrm{TLS_S} = \mathrm{LS_S} + n^{-1}2\mathrm{tr}(\hat{\mathbf{V}}_{\mathbf{s}}^{-1}\hat{\boldsymbol{\Lambda}}^{(1)}\hat{\boldsymbol{\Gamma}}^{(2)})$$
$$= \mathrm{LS_S} + n^{-1}2\mathrm{tr}\{(\hat{\boldsymbol{\Delta}}'\hat{\mathbf{V}}_{\mathbf{s}}^{-1}\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\mathbf{V}}_{\mathbf{s}}^{-1}\hat{\boldsymbol{\Gamma}}^{(2)}\hat{\mathbf{V}}_{\mathbf{s}}^{-1}\hat{\boldsymbol{\Delta}}\}$$

($\mathrm{ALS_S}$ and $\mathrm{CALS_S}$ are not defined), which is valid only when $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$.

Under possible non-normality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$,

$$\mathrm{E}_g(\mathrm{TLS_S}) - \mathrm{ECV_{SLS}} = O(n^{-2})$$

S2.4 $\mathrm{ALS_{ADFG}}$ and $\mathrm{CALS_{ADFG}}$ when $\hat{\mathbf{W}}_{\mathbf{s}} = \hat{\boldsymbol{\Gamma}}^{(2)} = n\,\widehat{\mathrm{acov}_{\mathrm{ADF}}(\mathbf{s})}$ by ADF-GLS for covariance structures

Recall that

$$\{n\,\widehat{\mathrm{acov}_{\mathrm{ADF}}(\mathbf{s})}\}_{ab,cd} = s_{abcd} - s_{ab}s_{cd}\ \ (p \ge a \ge b \ge 1;\ p \ge c \ge d \ge 1)$$

In this subsection, $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$ is assumed.

S2.4.1 Definition

Recall that

$$\mathrm{LS_{ADFG}} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'\hat{\boldsymbol{\Gamma}}^{(2)-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})\ \text{ and}$$
$$\mathrm{ELS_{ADFG}} = \mathrm{E}_g^{(\mathbf{s})}(\mathrm{LS_{ADFG}})$$

Define $\mathrm{ECV_{AGLS}} \equiv \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'\hat{\boldsymbol{\Gamma}}_{\mathbf{t}}^{(2)-1}(\mathbf{t} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})\}$.

S2.4.2 Bias of $\mathrm{LS_{ADFG}}$

Note that

$$\mathrm{ELS_{ADFG}} - \mathrm{ECV_{AGLS}}$$
$$= (\mathrm{ELS_{ADFG}} - \mathrm{EPLS_{ADFG}}) - (\mathrm{ECV_{AGLS}} - \mathrm{EPLS_{ADFG}}),$$

where the first term was given by (s1.4.4). Using the result before (s1.4.3), the reversed second term on the right-hand side of the above equation is

$$
\begin{aligned}
&\mathrm{ECV}_{\mathrm{AGLS}} - \mathrm{EPLS}_{\mathrm{ADFG}} \\
&= \mathrm{E}_g^{(\mathbf{t})}\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'(\hat{\boldsymbol{\Gamma}}_{\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_0^{(2)-1})(\mathbf{t}-\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})\} \\
&= [\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_0^{(2)-1})(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-2\mathrm{E}_g^{(\mathbf{t})}\{(\mathbf{t}-\boldsymbol{\sigma}_{\mathrm{T}})'(\hat{\boldsymbol{\Gamma}}_{\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_0^{(2)-1})\}_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{s})}(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}} - \boldsymbol{\sigma}_{\mathrm{T}})_{\to O(n^{-1})} \\
&+\mathrm{E}_g^{(\mathbf{t})}(\hat{\boldsymbol{\Gamma}}_{\mathbf{t}}^{(2)-1} - \boldsymbol{\Gamma}_0^{(2)-1})\}_{\to O(n^{-1})}\mathrm{E}_g^{(\mathbf{s})}\{(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}} - \boldsymbol{\sigma}_{\mathrm{T}})(\hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}} - \boldsymbol{\sigma}_{\mathrm{T}})'\}_{\to O(n^{-1})}]_{O(n^{-2})} \\
&+O(n^{-3}) \\
&= n^{-2}[n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-2n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{M}_{\mathrm{ADF}}^{(1)}\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)}) \\
&+\mathrm{tr}\{n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)})_{\to O(n^{-1})}\}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}'] + O(n^{-3})
\end{aligned}
\qquad \text{(s2.4.1)}
$$

(note that $\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_{\mathrm{ADF}}^{(1)}) = O(n^{-1})$ in the last result is due to e.g., $\mathrm{E}_g^{(\mathbf{s})}(s_{ancd} - \sigma_{\mathrm{T}abcd}) = O(n^{-1})$).

From (s1.4.4) and (s2.4.1),

$$
\begin{aligned}
&\mathrm{ELS}_{\mathrm{ADFG}} - \mathrm{ECV}_{\mathrm{AGLS}} \\
&= n^{-1}(-2q) + n^{-2}\underset{(\mathrm{A})}{[}\ 2\mathrm{tr}(\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(1)}\mathbf{K}_{(4)}) - 2\mathrm{tr}(\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(2)}\boldsymbol{\Gamma}_0^{(3)}) \\
&\qquad - 6\mathrm{tr}[\boldsymbol{\Gamma}_0^{(2)-1}\boldsymbol{\Lambda}_0^{(3)}\{\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)}) \otimes \boldsymbol{\Gamma}_0^{(2)}\}] \\
&+n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{I}_{(p^*)} - \boldsymbol{\Lambda}_0^{(1)})'(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)}) \\
&\qquad\qquad \times(\mathbf{I}_{(p^*)} - \boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-2n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'\boldsymbol{\Lambda}_0^{(2)}{}'\mathbf{M}_{\mathrm{ADF}}^{(1)}(\mathbf{I}_{(p^*)} - \boldsymbol{\Lambda}_0^{(1)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-n^2\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)})(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&+2n\mathrm{E}_g^{(\mathbf{s})}\{(\mathbf{s}-\boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{M}_{\mathrm{ADF}}^{(1)}\}_{\to O(n^{-1})}\boldsymbol{\Lambda}_0^{(2)}\mathrm{vec}(\boldsymbol{\Gamma}_0^{(2)}) \\
&-\mathrm{tr}\{n\mathrm{E}_g^{(\mathbf{s})}(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)})_{\to O(n^{-1})}\}\boldsymbol{\Lambda}_0^{(1)}\boldsymbol{\Gamma}_0^{(2)}\boldsymbol{\Lambda}_0^{(1)}{}']\underset{(\mathrm{A})}{]} + O(n^{-3}).
\end{aligned}
$$

which holds under possible non-normality. Under normality, $\mathbf{K}_{(4)} = \mathbf{O}$ and $\mathbf{\Gamma}_0^{(2)} = \mathbf{\Gamma}_{\mathrm{NT}}^{(2)}$.

S2.4.3 Bias correction of $\mathrm{LS}_{\mathrm{ADFG}}$

Recall that $\mathrm{LS}_{\mathrm{ADFG}} = (\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})'\hat{\mathbf{\Gamma}}^{(2)-1}(\mathbf{s} - \hat{\boldsymbol{\sigma}}_{\mathrm{AGLS}})$ and $\mathrm{ALS}_{\mathrm{ADFG}} = \mathrm{LS}_{\mathrm{ADFG}} + n^{-1}2q$.

Define

$$
\begin{aligned}
&\mathrm{CALS}_{\text{CV-ADFG}} = \mathrm{LS}_{\mathrm{ADFG}} + n^{-1}2q \\
&-n^{-2} \underset{(\mathrm{A})}{[} \; 2\mathrm{tr}(\hat{\mathbf{\Gamma}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(1)}\hat{\mathbf{K}}_{(4)}) - 2\mathrm{tr}(\hat{\mathbf{\Gamma}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(2)}\hat{\mathbf{\Gamma}}^{(3)}) \\
&\quad - 6\mathrm{tr}[\hat{\mathbf{\Gamma}}^{(2)-1}\hat{\mathbf{\Lambda}}^{(3)}\{\mathrm{vec}(\hat{\mathbf{\Gamma}}^{(2)}) \otimes \hat{\mathbf{\Gamma}}^{(2)}\}] \\
&+n^2 \widehat{\mathrm{E}_g^{(\mathbf{s})}}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})'(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)}) \\
&\qquad\qquad \times (\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-2n^2 \widehat{\mathrm{E}_g^{(\mathbf{s})}}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>}{}'\mathbf{\Lambda}_0^{(2)}{}'\mathbf{M}_{\mathrm{ADF}}^{(1)}(\mathbf{I}_{(p^*)} - \mathbf{\Lambda}_0^{(1)})(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&-n^2 \widehat{\mathrm{E}_g^{(\mathbf{s})}}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)})(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})\}_{\to O(n^{-2})} \\
&+2n\widehat{\mathrm{E}_g^{(\mathbf{s})}}\{(\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})'\mathbf{M}_{\mathrm{ADF}}^{(1)}\}_{\to O(n^{-1})}\mathbf{\Lambda}_0^{(2)}\mathrm{vec}(\mathbf{\Gamma}_0^{(2)}) \\
&-\mathrm{tr}\{n\widehat{\mathrm{E}_g^{(\mathbf{s})}}(\mathbf{M}_{\mathrm{ADF}}^{(1)} + \mathbf{M}_{\mathrm{ADF}}^{(2)})_{\to O(n^{-1})}\}\mathbf{\Lambda}_0^{(1)}\mathbf{\Gamma}_0^{(2)}\mathbf{\Lambda}_0^{(1)}{}'] \underset{(\mathrm{A})}{]}.
\end{aligned}
$$

$\mathrm{ALS}_{\mathrm{ADFG}}$ and $\mathrm{CALS}_{\mathrm{ADFG}}$ are valid only when $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$. Under possible nonnormality and $\boldsymbol{\sigma}_{\mathrm{T}} = \boldsymbol{\sigma}_0$

$\mathrm{E}_g(\mathrm{ALS}_{\mathrm{ADFG}}) - \mathrm{ECV}_{\mathrm{AGLS}} = O(n^{-2})$

and $\mathrm{E}_g(\mathrm{CALS}_{\mathrm{ADFG}}) - \mathrm{ECV}_{\mathrm{AGLS}} = O(n^{-3})$.

S2.5 $\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{ADFG}}$ by ADF-GLS using $\hat{\mathbf{\Gamma}}_{\boldsymbol{\rho}}^{(2)} = n\,\widehat{\mathrm{acov}_{\mathrm{ADF}}(\mathbf{r})}$ for correlation structures

S2.5.1 Definition

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}} = (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}}^{(2)-1}(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})$.

Define $\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{AGLS}} = \mathrm{E}_g^{(\mathbf{r}^*)}\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho},\mathbf{r}^*}^{(2)-1}(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})\}$. In this subsection $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$ is assumed.

S2.5.2 Bias of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}}$

Define $\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{ADFG}} \equiv \mathrm{E}_g^{(\mathbf{r})}(\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}})$. Then,

$$\begin{aligned}
&\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{ADFG}} - \mathrm{ECV}_{\boldsymbol{\rho}\mathrm{AGLS}} \\
&= (\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{ADFG}} - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{ADFG}}) - (\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{AGLS}} - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{ADFG}}) \\
&= \mathrm{ELS}_{\boldsymbol{\rho}\mathrm{ADFG}} - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{ADFG}} \\
&\quad - \mathrm{E}_g^{(\mathbf{r}^*)}\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})'(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho},\mathbf{r}^*}^{(2)-1} - \boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)-1})(\mathbf{r}^* - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})\} \\
&= \mathrm{ELS}_{\boldsymbol{\rho}\mathrm{ADFG}} - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{ADFG}} + O(n^{-2}) \\
&= -n^{-1}2q + O(n^{-2})
\end{aligned}$$

(see Subsection S1.5.2).

S2.5.3 Bias correction of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}}$

Recalling that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}} = (\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}}^{(2)-1}(\mathbf{r} - \hat{\boldsymbol{\rho}}_{\mathrm{AGLS}})$ and $\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{ADFG}} = \mathrm{LS}_{\boldsymbol{\rho}\mathrm{ADFG}} + n^{-1}2q$ ($\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{ADFG}}$ is unnecessary), which is valid only when $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$.

Under possible non-normality and $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$,
$\mathrm{E}_g(\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{ADFG}}) - \mathrm{ECV}_{\boldsymbol{\rho}\mathrm{AGLS}} = O(n^{-2})$, which is the same as that in Subsection 1.5.3 up to this order.

S2.6 $\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{NTG}}$ by NT-GLS using $\hat{\boldsymbol{\Gamma}}_{\boldsymbol{\rho}\mathrm{NT}}^{(2)} = n\,\widehat{\mathrm{acov}_{\mathrm{NT}}(\mathbf{r})}$ for correlation structures

S2.6.1 Definition

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}=(\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}(\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})$ with $\hat{\boldsymbol{\Gamma}}^{(2)}_{\boldsymbol{\rho}\mathrm{NT}}=\hat{\boldsymbol{\Gamma}}^{(2)}_{\boldsymbol{\rho}\mathrm{NT},\mathbf{r}}$. Define $\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{NGLS}}=\mathrm{E}^{(\mathbf{r}^*)}_g\mathrm{E}^{(\mathbf{r})}_g\{(\mathbf{r}^*-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT},\mathbf{r}^*}(\mathbf{r}^*-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})\}$. In this subsection $\boldsymbol{\rho}_{\mathrm{T}}=\boldsymbol{\rho}_0$ is assumed.

S2.6.2 Bias of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}$

Define $\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{NTG}}=\mathrm{E}^{(\mathbf{r})}_g(\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}})$. Then,

$$\begin{aligned}
&\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{NTG}}-\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{NGLS}}\\
&=(\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{NTG}}-\mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}})-(\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{NGLS}}-\mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}})\\
&=\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{NTG}}-\mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}}\\
&\quad-\mathrm{E}^{(\mathbf{r}^*)}_g\mathrm{E}^{(\mathbf{r})}_g\{(\mathbf{r}^*-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'(\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT},\mathbf{r}^*}-\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}})(\mathbf{r}^*-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})\}\\
&=\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{NTG}}-\mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{NTG}}+O(n^{-2})\\
&=-n^{-1}2\mathrm{tr}\{(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}{}'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}{}'\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Gamma}^{(2)}_{\boldsymbol{\rho}}\boldsymbol{\Gamma}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}\}+O(n^{-2}),
\end{aligned}$$

which becomes $-n^{-1}2q$ under normality.

S2.6.3 Bias correction of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}$

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}=(\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}(\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})$, $\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{NTG}}=\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}+n^{-1}2q$ and $\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{NTG}}=\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}+n^{-1}2\mathrm{tr}\{(\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}{}'\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}})^{-1}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}{}'\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\hat{\boldsymbol{\Gamma}}^{(2)}\hat{\boldsymbol{\Gamma}}^{(2)-1}_{\boldsymbol{\rho}\mathrm{NT}}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}\}$, which are valid only when $\boldsymbol{\rho}_{\mathrm{T}}=\boldsymbol{\rho}_0$.

Under normality and $\boldsymbol{\rho}_{\mathrm{T}}=\boldsymbol{\rho}_0$,

$\mathrm{E}_f(\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{NTG}})-\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{NGLS}}=O(n^{-2})$

and $\mathrm{E}_f(\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{NTG}})-\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{NGLS}}=O(n^{-2})$.

Under non-normality and $\boldsymbol{\rho}_{\mathrm{T}}=\boldsymbol{\rho}_0$,

$\mathrm{E}_g(\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{NTG}})-\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{NGLS}}=O(n^{-2})$, which is the same as that in Subsection

1.6.3 up to this order.

S2.7 $\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{U}}$ by ULS for correlation structures

S2.7.1 Definition

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}} \equiv (\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'(\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}}) = (1/2)\mathrm{tr}\{(\mathbf{R}-\hat{\mathbf{P}}_{\mathrm{ULS}})^2\}$ ($\mathrm{Diag}(\hat{\mathbf{P}}_{\mathrm{ULS}}) = \mathbf{I}_{(p)}$ is assumed).

Define $\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{ULS}} \equiv \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{U}} = \mathrm{E}_g^{(\mathbf{r}^*)}\mathrm{E}_g^{(\mathbf{r})}\{(\mathbf{r}^*-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})'(\mathbf{r}^*-\hat{\boldsymbol{\rho}}_{\mathrm{NGLS}})\}$. In this subsection $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$ is assumed.

S2.7.2 Bias of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{NTG}}$

Define $\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{U}} \equiv \mathrm{E}_g^{(\mathbf{r})}(\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}})$. Then,

$$\mathrm{ELS}_{\boldsymbol{\rho}\mathrm{U}} - \mathrm{ECV}_{\boldsymbol{\rho}\mathrm{ULS}} = \mathrm{ELS}_{\boldsymbol{\rho}\mathrm{U}} - \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{U}}$$
$$= -n^{-1}2\mathrm{tr}\{(\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Delta}_{\boldsymbol{\rho}0})^{-1}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}'\boldsymbol{\Gamma}_{\boldsymbol{\rho}}^{(2)}\boldsymbol{\Delta}_{\boldsymbol{\rho}0}\} + O(n^{-2}).$$

S2.7.3 Bias correction of $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}}$

Recall that $\mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}} = (\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{ULS}})'(\mathbf{r}-\hat{\boldsymbol{\rho}}_{\mathrm{ULS}})$,

and $\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{U}} = \mathrm{LS}_{\boldsymbol{\rho}\mathrm{U}} + n^{-1}2\mathrm{tr}\{(\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}'\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}})^{-1}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}'\hat{\boldsymbol{\Gamma}}^{(2)}\hat{\boldsymbol{\Delta}}_{\boldsymbol{\rho}}\}$.

($\mathrm{ALS}_{\boldsymbol{\rho}\mathrm{U}}$ and $\mathrm{CALS}_{\boldsymbol{\rho}\mathrm{U}}$ are not defined), which is valid only when $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$.

Under possible non-normality and $\boldsymbol{\rho}_{\mathrm{T}} = \boldsymbol{\rho}_0$,

$\mathrm{E}_g(\mathrm{TLS}_{\boldsymbol{\rho}\mathrm{U}}) - \mathrm{ECV}_{\boldsymbol{\rho}\mathrm{ULS}} = O(n^{-2})$, which is exactly the same as that in Subsection 1.7.3 since by definition $\mathrm{ECV}_{\boldsymbol{\rho}\mathrm{ULS}} = \mathrm{EPLS}_{\boldsymbol{\rho}\mathrm{U}}$ in this case.

S3. Miscellaneous results

S3.1 Explicit expressions of the elements of $\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}^{-1}$

It is known that

$$\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}_{ab,cd}=\{2\mathbf{D}_p^{+}(\boldsymbol{\Sigma}_{\mathrm{T}}\otimes\boldsymbol{\Sigma}_{\mathrm{T}})\mathbf{D}_p^{+}{}'\}_{ab,cd}=\sigma_{\mathrm{T}ac}\sigma_{\mathrm{T}bd}+\sigma_{\mathrm{T}ad}\sigma_{\mathrm{T}bc}$$
$(1\le a\le b\le p;\ 1\le c\le d\le p)$.

We derive the elements associated with $X_a,\ X_b,\ X_c$ and X_d $(1\le a\le b\le p;\ 1\le c\le d\le p)$ for $[\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}^{-1}]_{ab,cd}=\{(1/2)\mathbf{D}_p{}'(\boldsymbol{\Sigma}_{\mathrm{T}}^{-1}\otimes\boldsymbol{\Sigma}_{\mathrm{T}}^{-1})\mathbf{D}_p\}_{ab,cd}$. The 3×3 asymmetric matrix for the elements using double subscript notation is

$$[\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}^{-1}]_{(aa,ba,bb;\ cc,dc,dd)}$$
$$=\frac{1}{2}\begin{pmatrix}1&0&0&0\\0&1&1&0\\0&0&0&1\end{pmatrix}\begin{bmatrix}\sigma_{\mathrm{T}}^{ac}\begin{pmatrix}\sigma_{\mathrm{T}}^{ac}&\sigma_{\mathrm{T}}^{ad}\\\sigma_{\mathrm{T}}^{bc}&\sigma_{\mathrm{T}}^{bd}\end{pmatrix}&\sigma_{\mathrm{T}}^{ad}\begin{pmatrix}\sigma_{\mathrm{T}}^{ac}&\sigma_{\mathrm{T}}^{ad}\\\sigma_{\mathrm{T}}^{bc}&\sigma_{\mathrm{T}}^{bd}\end{pmatrix}\\\sigma_{\mathrm{T}}^{bc}\begin{pmatrix}\sigma_{\mathrm{T}}^{ac}&\sigma_{\mathrm{T}}^{ad}\\\sigma_{\mathrm{T}}^{bc}&\sigma_{\mathrm{T}}^{bd}\end{pmatrix}&\sigma_{\mathrm{T}}^{bd}\begin{pmatrix}\sigma_{\mathrm{T}}^{ac}&\sigma_{\mathrm{T}}^{ad}\\\sigma_{\mathrm{T}}^{bc}&\sigma_{\mathrm{T}}^{bd}\end{pmatrix}\end{bmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&1&0\\0&0&1\end{pmatrix}$$
$$=\frac{1}{2}\begin{pmatrix}(\sigma_{\mathrm{T}}^{ac})^2&2\sigma_{\mathrm{T}}^{ac}\sigma_{\mathrm{T}}^{ad}&(\sigma_{\mathrm{T}}^{ad})^2\\2\sigma_{\mathrm{T}}^{ac}\sigma_{\mathrm{T}}^{bc}&2(\sigma_{\mathrm{T}}^{ac}\sigma_{\mathrm{T}}^{bd}+\sigma_{\mathrm{T}}^{ad}\sigma_{\mathrm{T}}^{bc})&2\sigma_{\mathrm{T}}^{ad}\sigma_{\mathrm{T}}^{bd}\\(\sigma_{\mathrm{T}}^{bc})^2&2\sigma_{\mathrm{T}}^{bc}\sigma_{\mathrm{T}}^{bd}&(\sigma_{\mathrm{T}}^{bd})^2\end{pmatrix}.$$

From this expression, we have

$$[\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}^{-1}]_{ab,cd}=(1/4)(2-\delta_{ab})(2-\delta_{cd})(\sigma_{\mathrm{T}}^{ac}\sigma_{\mathrm{T}}^{bd}+\sigma_{\mathrm{T}}^{ad}\sigma_{\mathrm{T}}^{bc})$$
$(1\le a\le b\le p;\ 1\le c\le d\le p)$.

The result is confirmed as follows.

$$[\{n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})\}^{-1}n\,\mathrm{cov}_{\mathrm{NT}}(\mathbf{s})]_{ab,ef}\,(1\le a\le b\le p;\ 1\le e\le f\le p)$$
$$=\sum_{c\ge d}(1/4)(2-\delta_{ab})(2-\delta_{cd})(\sigma_{\mathrm{T}}^{ac}\sigma_{\mathrm{T}}^{bd}+\sigma_{\mathrm{T}}^{ad}\sigma_{\mathrm{T}}^{bc})$$
$$\times(\sigma_{\mathrm{T}ce}\sigma_{\mathrm{T}df}+\sigma_{\mathrm{T}cf}\sigma_{\mathrm{T}de})$$
$$=\sum_{c=1}^{p}\sum_{d=1}^{p}\frac{2-\delta_{ab}}{2}\sigma_{\mathrm{T}}^{ac}\sigma_{\mathrm{T}}^{bd}(\sigma_{\mathrm{T}ce}\sigma_{\mathrm{T}df}+\sigma_{\mathrm{T}cf}\sigma_{\mathrm{T}de})\qquad\text{(s3.1.1)}$$
$$=\frac{2-\delta_{ab}}{2}(\delta_{ae}\delta_{bf}+\delta_{af}\delta_{be}).$$

When $a > b$ and $e > f$, (s3.1.1) gives $\delta_{ae}\delta_{bf}$,

when $a = b$ and $e > f$, (s3.1.1) gives $\delta_{ae}\delta_{bf}(=0)$,

when $a > b$ and $e = f$, (s3.1.1) gives $\delta_{ae}\delta_{bf}(=0)$

and when $a = b$ and $e = f$, (s3.1.1) gives $\delta_{ae}\delta_{bf}$.

This shows that $[\cdot]_{ab,ef}$ is $[\mathbf{I}_{(p^*)}]_{ab,ef}$.

S3.2 minimization of $F \equiv (1/2)\mathrm{tr}[\{\boldsymbol{\Sigma}^{-1}(\mathbf{S}-\boldsymbol{\Sigma})\}^2]$
$= (1/2)\mathrm{tr}\{(\mathbf{I}_{(p)} - \boldsymbol{\Sigma}^{-1}\mathbf{S})^2\}$ **with respect to** $\boldsymbol{\theta}$ **in** $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$

$$\frac{\partial F}{\partial \theta_i} = -\mathrm{tr}\left\{(\mathbf{I}_{(p)} - \boldsymbol{\Sigma}^{-1}\mathbf{S})\frac{\partial \boldsymbol{\Sigma}^{-1}\mathbf{S}}{\partial \theta_i}\right\} = \mathrm{tr}\left\{(\mathbf{I}_{(p)} - \boldsymbol{\Sigma}^{-1}\mathbf{S})\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\boldsymbol{\Sigma}^{-1}\mathbf{S}\right\}$$

$$= \mathrm{tr}\left\{\boldsymbol{\Sigma}^{-1}(\mathbf{S} - \mathbf{S}\boldsymbol{\Sigma}^{-1}\mathbf{S})\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right\},$$

$$\frac{\partial^2 F}{\partial \theta_i \partial \theta_j} \doteq \mathrm{tr}\left\{\boldsymbol{\Sigma}^{-1}\mathbf{S}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\boldsymbol{\Sigma}^{-1}\mathbf{S}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\right\} \tag{s3.2.1}$$

$$\doteq \mathrm{tr}\left\{\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\right\} \ (i, j = 1, \ldots, q). \tag{s3.2.2}$$

For an iterative computation, (s3.2.2) can be used. However, (s3.2.1) seems to give somewhat faster computation than that of (3.2.2).

References

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research, 24*, 445-455.

Girshick, M. A. (1939). On the sampling theory of roots of determinantal equations. *Annals of Mathematical Statistics, 10,* 203-224.

Holmquist, B. (1988). Moments and cumulants of the multivariate normal distribution. *Stochastic Analysis and Applications, 6*, 273-278.

Hsu, P. L. (1949). The limiting distribution of functions of sample means and application to testing hypotheses. In J. Neyman (Ed.), *Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability* (pp.359-402). Berkeley: University of California Press.

Isserlis, L. (1916). On certain probable errors and correlation coefficients of multiple

frequency distributions with skew regression. *Biometrika, 11*, 185-190
Kano, Y. (1997). Beyond third-order efficiency. *Sankhya, B, 59*, 179-197.
Magnus, J. R., & Neudecker, H. (1999). *Matrix differential calculus with applications in statistics and econometrics* (Rev. ed.). New York: Wiley.
Ogasawara, H. (2002). Concise formulas for the standard errors of component loading estimates. *Psychometrika*, *67*, 289-297.
Ogasawara, H. (2006). Asymptotic expansion of the sample correlation coefficient under nonnormality. *Computational Statistics and Data Analysis, 50,* 891-910.
Ogasawara, H. (2007). Higher-order estimation error in structural equation modeling. *Economic Review, Otaru University of Commerce*, *57* (4), 131-160. http://www.geocities.jp/hogasa07/, Permalink: http://hdl.handle.net/10252/268.
Ogasawara, H. (2008). Asymptotic expansions of the distribution of the estimator for the generalized partial correlation under nonnormality. *Behaviormetrika, 35,* 15-33.
Ogasawara, H. (2009). Asymptotic expansions in mean and covariance structure analysis. *Journal of Multivariate Analysis, 100*, 902-919.
Ogasawara, H. (2010). Supplement to the paper "Asymptotic expansions of the distributions of the chi-square statistic based on the asymptotically distribution-free theory in covariance structures". *Economic Review* (*Otaru University of Commerce*), *60* (4), 187-200. http://www.geocities.jp/hogasa07/, Permalink: http://hdl.handle.net/10252/4122.
Ogasawara, H. (2017). Expected predictive least squares for model selection in covariance structures. *Journal of Multivariate Analysis, 155,* 151-164.
Olkin, I., & Siotani, M. (1976). Asymptotic distribution of functions of a correlation matrix. In S. Ikeda et al. (Eds.) *Essays in provability and statistics: A volume in honor of Professor Junjiro Ogawa* (pp. 235-251). Tokyo: Shinko Tsusho.
Pearson, K., & Filon, L. N. G. (1898). Mathematical contributions to the theory of evolution.-IV. On the probable errors of frequency constants and on the influence of random selection on variation and correlation. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 191,* 229-311.
Steiger, J. H., & Hakstian, A. R. (1982). The asymptotic distribution of elements of a correlation matrix: Theory and application. *British Journal of Mathematical and Statistical Psychology, 35,* 208-215.
Steiger, J. H., & Hakstian, A. R. (1983). A historical note on the asymptotic distribution of correlations. *British Journal of Mathematical and Statistical Psychology, 36,* 157.