SUPPLEMENT TO
PRINCIPAL COMPONENTS MODELS FOR CORRELATION MATRICES

Robert J. Boik
Department of Mathematical Sciences
Montana State University — Bozeman
Bozeman, MT 59717-2400

Email: rjboik@math.montana.edu

November 7, 2008

Contents

1 Introduction

This supplement gives proofs, higher order derivatives, and other details concerning the parameterization correlation matrices as well as the distributions of the estimators. The format of this supplement closely follows that of the supplement to Boik (2002).

The notation in this supplement departs slightly from that in the manuscript “Principal Components Models for Correlation Matrices.” First, vectors and matrices are set in boldface type for easy recognition. More importantly, it is necessary to assemble the p variables into k sets, where the i^{th} set contains p_i variables and $\sum_{i=1}^{k} p_i = p$. The boldface symbol $\mathbf{p}$ will be used to denote the k-vector $\begin{pmatrix} p_1 & p_2 & \cdots & p_k \end{pmatrix}'$, and, to avoid confusion among symbols, the number of variables will be denoted by $\dot{p}$ rather than p.

Several typographical errors in Theorems 10 and 11 were corrected on 11/24/07 and on 11/07/08. These errors do not affect the Matlab programs or the numerical results in the 2002 article.

2 Proof of Theorem 1

Theorem 1. *Let* $\mathcal{P}_{\dot{p}}$ *be the space of all* $\dot{p} \times \dot{p}$ *permutation matrices. For* $\mathbf{Q} \in \mathcal{P}_{\dot{p}}$*, define* $\boldsymbol{\Psi}(\mathbf{Q})$ *as* $\boldsymbol{\Psi}(\mathbf{Q}) \overset{\text{def}}{=} \mathbf{Q}'\boldsymbol{\Psi}\mathbf{Q}$ *and define* $\operatorname{blk}(\mathbf{Q})$ *to be be the maximal number of non-empty blocks for which*

$$\boldsymbol{\Psi}(\mathbf{Q}) = \bigoplus_{i=1}^{\operatorname{blk}(\mathbf{Q})} [\boldsymbol{\Psi}(\mathbf{Q})]_{ii}$$

is satisfied, where $\oplus$ *is the direct sum operator (Searle, 1982, §10.6) and* $[\boldsymbol{\Psi}(\mathbf{Q})]_{ii}$ *is the* i^{th} *non-empty block on the main diagonal of* $\boldsymbol{\Psi}(\mathbf{Q})$*. Also, define* k *as*

$$k = \max_{\mathbf{Q} \in \mathcal{P}_{\dot{p}}} \operatorname{blk}(\mathbf{Q}).$$

Then, $\operatorname{rank}(\mathbf{W}_{\boldsymbol{\psi}}) = \dot{p} - k$ *and* $\mathbf{W}_{\boldsymbol{\psi}}$ *has full row-rank if and only if* $k = 1$*.*

Proof. First, it will be shown that $\operatorname{rank}(\mathbf{W}_{\boldsymbol{\psi}}) \le \dot{p} - k$*. Suppose that* $\boldsymbol{\Psi}(\mathbf{Q})$ *has a block diagonal structure with* k *blocks, where the* i^{th} *block has dimension* $p_i \times p_i$*. Write the* i^{th} *block as* $[\boldsymbol{\Psi}(\mathbf{Q})]_{ii} = \boldsymbol{\Gamma}^*_{ii}\boldsymbol{\Lambda}^*_{ii}\boldsymbol{\Gamma}^{*\prime}_{ii}$*, where* $\boldsymbol{\Gamma}^*_{ii} = [\mathbf{Q}'\boldsymbol{\Gamma}\mathbf{Q}^*]_{ii} \in \mathcal{O}_{p_i}$*,* $\mathbf{Q}^* \in \mathcal{P}_{\dot{p}}$*, and* $\boldsymbol{\Lambda}^*_{ii} = [\mathbf{Q}^{*\prime}\boldsymbol{\Lambda}\mathbf{Q}^*]_{ii}$*. Let* $\mathbf{c}_1, \ldots, \mathbf{c}_{k-1}$ *be a set of* $\dot{p}$*-vectors whose elements consist of coefficients of* $k-1$ *linearly independent contrasts among the* k *blocks. Specifically, define* $\mathbf{c}_j$ *as*

$$\mathbf{c}_j = \frac{1}{p_j}\mathbf{E}_{j,\mathbf{p}}\mathbf{1}_{p_j} - \frac{1}{p_k}\mathbf{E}_{k,\mathbf{p}}\mathbf{1}_{p_k},$$

where $\mathbf{E}_{j,\mathbf{p}}$ *is defined in* (4)*. It follows that* $\mathbf{c}_j'\mathbf{1}_{\dot{p}} = 0$ *and that* $\mathbf{L}\mathbf{Q}\mathbf{c}_j = \mathbf{C}_3\boldsymbol{\ell}_j$*, where* $\boldsymbol{\ell}_j = (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{Q}\mathbf{c}_j$*,* $\mathbf{L}$ *is defined in* (2)*, and* $\mathbf{C}_3$ *and* $\mathbf{K}$ *are defined in* (9)*. Furthermore,*

$$\begin{aligned}
\mathbf{W}_{\boldsymbol{\psi}}'\boldsymbol{\ell}_j &= \mathbf{A}_2'\left(\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})}\right)(\boldsymbol{\Lambda}\boldsymbol{\Gamma}' \otimes \boldsymbol{\Gamma}')\mathbf{L}\mathbf{Q}\mathbf{c}_j \\
&= \mathbf{A}_2'\left(\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})}\right)(\mathbf{Q}^*\mathbf{Q}^{*\prime} \otimes \mathbf{Q}^*\mathbf{Q}^{*\prime})(\boldsymbol{\Lambda}\boldsymbol{\Gamma}' \otimes \boldsymbol{\Gamma}')\mathbf{L}\mathbf{Q}\mathbf{c}_j \\
&= \mathbf{A}_2'(\mathbf{Q}^* \otimes \mathbf{Q}^*)\left(\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})}\right)(\mathbf{Q}^{*\prime}\boldsymbol{\Lambda}\boldsymbol{\Gamma}' \otimes \mathbf{Q}^{*\prime}\boldsymbol{\Gamma}')\operatorname{vec}(\mathbf{Q}\mathbf{D}_j\mathbf{Q}')
\end{aligned}$$

$$\text{where } \mathbf{D}_j = \operatorname{Diag}(\mathbf{c}_j) = \frac{1}{p_j}\mathbf{E}_{j,\mathbf{p}}\mathbf{E}_{j,\mathbf{p}}' - \frac{1}{p_k}\mathbf{E}_{k,\mathbf{p}}\mathbf{E}_{k,\mathbf{p}}',$$

$$\begin{aligned}
&= \mathbf{A}_2'(\mathbf{Q}^*\otimes\mathbf{Q}^*)\left(\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}\right)\operatorname{vec}\left(\mathbf{Q}^{*\prime}\boldsymbol{\Gamma}'\mathbf{Q}\mathbf{D}_i\mathbf{Q}'\boldsymbol{\Gamma}\mathbf{Q}^*\mathbf{Q}^{*\prime}\boldsymbol{\Lambda}\mathbf{Q}^*\right)\\
&= \mathbf{A}_2'(\mathbf{Q}^*\otimes\mathbf{Q}^*)\left(\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}\right)\operatorname{vec}\left(\frac{1}{p_j}\mathbf{E}_{j,\mathbf{p}}\boldsymbol{\Gamma}^{*\prime}_{jj}\boldsymbol{\Gamma}^*_{jj}\boldsymbol{\Lambda}^*_{jj}\mathbf{E}'_{j,\mathbf{p}}-\frac{1}{p_k}\mathbf{E}_{k,\mathbf{p}}\boldsymbol{\Gamma}^{*\prime}_{kk}\boldsymbol{\Gamma}^*_{kk}\boldsymbol{\Lambda}^*_{kk}\mathbf{E}'_{k,\mathbf{p}}\right)\\
&= \mathbf{A}_2'(\mathbf{Q}^*\otimes\mathbf{Q}^*)\left(\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}\right)\operatorname{vec}\left(\frac{1}{p_j}\mathbf{E}_{j,\mathbf{p}}\boldsymbol{\Lambda}^*_{jj}\mathbf{E}'_{j,\mathbf{p}}-\frac{1}{p_k}\mathbf{E}_{k,\mathbf{p}}\boldsymbol{\Lambda}^*_{kk}\mathbf{E}'_{k,\mathbf{p}}\right)=\mathbf{0}
\end{aligned}$$

because $\mathbf{E}_{j,\mathbf{p}}\boldsymbol{\Lambda}^*_{jj}\mathbf{E}'_{j,\mathbf{p}}$ *and* $\mathbf{E}_{k,\mathbf{p}}\boldsymbol{\Lambda}^*_{kk}\mathbf{E}'_{k,\mathbf{p}}$ *are symmetric and* $\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}$ *projects onto the space of skew symmetric matrices along the space of symmetric matrices. Linear independence of* $\mathbf{c}_1,\ldots,\mathbf{c}_{k-1}$ *implies that* $\operatorname{rank}(\mathbf{W}_{\boldsymbol{\psi}})\le\dot p-k$*.*

Second, it will be shown that $\operatorname{rank}(\mathbf{W}_{\boldsymbol{\psi}})<\dot p-k$ *leads to a contradiction. If* $\operatorname{rank}(\mathbf{W}_{\boldsymbol{\psi}})<\dot p-k$*, then there exists a* $\dot p$*-vector* $\mathbf{c}_k$ *such that* $\mathbf{c}_k'\mathbf{1}_{\dot p}=0$, $\operatorname{rank}\begin{pmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_k\end{pmatrix}=k$*, and* $\mathbf{W}_{\boldsymbol{\psi}}'\boldsymbol{\ell}_k=\mathbf{0}$*, where* $\boldsymbol{\ell}_k=(\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{c}_k$*. Denote the null space (kernel) of a matrix* $\mathbf{M}$ *by* $\mathcal{N}(\mathbf{M})$*. Define* $\mathbf{J}_{\mathbf{m}}$ *as*

$$\mathbf{J}_{\mathbf{m}}=\bigoplus_{i=1}^{\dim(\mathbf{m})}\mathbf{J}_{m_i},$$

where $\mathbf{m}$ *is the vector of eigenvalue multiplicities defined in §2.3 of the article and* $\mathbf{J}_{m_i}$ *is an* $m_i\times m_i$ *matrix of ones. Let* $\omega_{\mathbf{m}}$ *be the set*

$$\omega_{\mathbf{m}}=\left\{(s,t);t>s,\mathbf{e}_s^{\dot p\prime}\mathbf{J}_{\mathbf{m}}\mathbf{e}_t^{\dot p}=1\right\}.$$

That is, $\omega_{\mathbf{m}}$ *is the set of row and column indices that correspond to the structural zeros in* $\mathbf{G}$*. The set* $\omega_{\mathbf{m}}$ *contains* $q_{\mathbf{m}}=(\mathbf{m}'\mathbf{m}-\dot p)/2$ *index pairs. For convenience, label the pairs as* (s_ℓ,t_ℓ) *for* $\ell=1,\ldots,q_{\mathbf{m}}$*. It can be shown that a set of vectors that span the* $\dot p(\dot p+1)/2+q_{\mathbf{m}}$ *dimensional space* $\mathcal{N}\left[\mathbf{A}_2'\left(\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}\right)(\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\right]$ *is given by the columns of*

$$\mathbf{X}=(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma})\begin{pmatrix}\mathbf{N}_{\dot p} & \mathbf{T}\end{pmatrix},\ \textit{where}\ \mathbf{T}=\sum_{\ell=1}^{q_{\mathbf{m}}}(\mathbf{e}_{t_\ell}^{\dot p}\otimes\mathbf{e}_{s_\ell}^{\dot p})\mathbf{e}_\ell^{\dot p\prime},$$

and $\mathbf{N}_{\dot p}$ *is defined in* (22)*. Let* $\mathbf{P}_x=\mathbf{X}\boldsymbol{\Omega}^{-1}\left(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}\right)^-\mathbf{X}'\boldsymbol{\Omega}^{-1}$*, where* $\boldsymbol{\Omega}$ *is a* $\dot p^2\times\dot p^2$ *positive definite matrix. Note that* $\mathbf{P}_x$ *is the linear operator that projects onto* $\mathcal{R}(\mathbf{X})=\mathcal{N}\left[\mathbf{A}_2'\left(\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}\right)(\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\right]$ *along* $\mathcal{N}(\mathbf{X}'\boldsymbol{\Omega}^{-1})$*. Therefore,*

$$\mathbf{W}_{\boldsymbol{\psi}}'\boldsymbol{\ell}_k=\mathbf{0}\iff\mathbf{P}_x\mathbf{C}_3\boldsymbol{\ell}_k=\mathbf{C}_3\boldsymbol{\ell}_k.$$

Set $\boldsymbol{\Omega}^{-1}=(\boldsymbol{\Psi}^2\otimes\mathbf{I}_{\dot p})$*. It can be shown that one generalized inverse of* $\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}$ *is*

$$\left(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}\right)^-=\begin{pmatrix}\mathbf{N}_{\dot p}+2\,\mathbf{N}_{\dot p}\,\mathbf{T}\mathbf{T}'\,\mathbf{N}_{\dot p} & -2\,\mathbf{N}_{\dot p}\,\mathbf{T}\\ -2\mathbf{T}'\,\mathbf{N}_{\dot p} & 2\mathbf{I}_{q_{\mathbf{m}}}\end{pmatrix}.$$

Accordingly,

$$\mathbf{P}_x=(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma})\left[\mathbf{N}_{\dot p}+\frac{1}{2}\left(\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}\right)\mathbf{T}\mathbf{T}'\left(\mathbf{I}_{\dot p^2}-\mathbf{I}_{(\dot p,\dot p)}\right)\right](\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')$$

and

$$\begin{aligned}
\mathbf{P}_x\mathbf{C}_3\boldsymbol{\ell}_k &= \mathbf{C}_3\boldsymbol{\ell}_k\iff\\
\mathbf{P}_x\mathbf{C}_3\boldsymbol{\ell}_k &= (\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma})\,\mathbf{N}_{\dot p}(\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\mathbf{C}_3\boldsymbol{\ell}_k \qquad (33)
\end{aligned}$$

$$+\frac{1}{2}(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma})\left(\mathbf{I}_{\dot{p}^2}-\mathbf{I}_{(\dot{p},\dot{p})}\right)\mathbf{T}\mathbf{T}'\left(\mathbf{I}_{\dot{p}^2}-\mathbf{I}_{(\dot{p},\dot{p})}\right)(\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\mathbf{C}_3\boldsymbol{\ell}_k.$$

The second term on the right-hand-side of (33) *is zero because* $\mathbf{C}_3\boldsymbol{\ell}_k=\mathbf{L}\mathbf{c}_k$ *and*

$$\begin{aligned}
&(\boldsymbol{\Lambda}^{-1}\otimes\mathbf{I}_{\dot{p}})\left(\mathbf{I}_{\dot{p}^2}-\mathbf{I}_{(\dot{p},\dot{p})}\right)\mathbf{T}\mathbf{T}'\left(\mathbf{I}_{\dot{p}^2}-\mathbf{I}_{(\dot{p},\dot{p})}\right)(\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\mathbf{L}\mathbf{c}_k\\
&=\sum_{i=1}^{\dot{p}}\sum_{(s,t)\in\omega_{\mathbf{m}}}(\boldsymbol{\Lambda}^{-1}\otimes\mathbf{I}_{\dot{p}})\Big[\left(\mathbf{e}_t^{\dot{p}}\mathbf{e}_t^{\dot{p}\prime}\otimes\mathbf{e}_s^{\dot{p}}\mathbf{e}_s^{\dot{p}\prime}\right)-\left(\mathbf{e}_s^{\dot{p}}\mathbf{e}_t^{\dot{p}\prime}\otimes\mathbf{e}_t^{\dot{p}}\mathbf{e}_s^{\dot{p}\prime}\right)\\
&\quad-\left(\mathbf{e}_t^{\dot{p}}\mathbf{e}_s^{\dot{p}\prime}\otimes\mathbf{e}_s^{\dot{p}}\mathbf{e}_t^{\dot{p}\prime}\right)+\left(\mathbf{e}_s^{\dot{p}}\mathbf{e}_s^{\dot{p}\prime}\otimes\mathbf{e}_t^{\dot{p}}\mathbf{e}_t^{\dot{p}\prime}\right)\Big](\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\left(\mathbf{e}_i^{\dot{p}}\otimes\mathbf{e}_i^{\dot{p}}\right)\mathbf{e}_i^{\dot{p}\prime}\mathbf{c}_k\\
&=\sum_{i=1}^{\dot{p}}\sum_{(s,t)\in\omega_{\mathbf{m}}}\Big[\left(\mathbf{e}_t^{\dot{p}}\lambda_t^{-1}\lambda_t\otimes\mathbf{e}_s^{\dot{p}}\lambda_s^{-1}\lambda_s\right)\gamma_{ti}\gamma_{si}-\left(\mathbf{e}_s^{\dot{p}}\lambda_s^{-1}\lambda_t\otimes\mathbf{e}_t^{\dot{p}}\lambda_t^{-1}\lambda_s\right)\gamma_{ti}\gamma_{si}\\
&\quad-\left(\mathbf{e}_t^{\dot{p}}\lambda_t^{-1}\lambda_s\otimes\mathbf{e}_s^{\dot{p}}\lambda_s^{-1}\lambda_t\right)\gamma_{ti}\gamma_{si}+\left(\mathbf{e}_s^{\dot{p}}\lambda_s^{-1}\lambda_s\otimes\mathbf{e}_t^{\dot{p}}\lambda_t^{-1}\lambda_t\right)\gamma_{ti}\gamma_{si}\Big]\mathbf{e}_i^{\dot{p}\prime}\mathbf{c}_k\\
&=\sum_{i=1}^{\dot{p}}\sum_{(s,t)\in\omega_{\mathbf{m}}}\Big[\left(\mathbf{e}_t^{\dot{p}}\otimes\mathbf{e}_s^{\dot{p}}\right)-\left(\mathbf{e}_s^{\dot{p}}\otimes\mathbf{e}_t^{\dot{p}}\right)-\left(\mathbf{e}_t^{\dot{p}}\otimes\mathbf{e}_s^{\dot{p}}\right)+\left(\mathbf{e}_s^{\dot{p}}\otimes\mathbf{e}_t^{\dot{p}}\right)\Big]\gamma_{ti}\gamma_{si}\mathbf{e}_i^{\dot{p}\prime}\mathbf{c}_k=\mathbf{0}.
\end{aligned}$$

From the remaining terms of (33), *it follows that*

$$\begin{aligned}
&\mathbf{P}_x\mathbf{L}\mathbf{c}_k=\mathbf{L}\mathbf{c}_k\\
&\iff(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma})\,\mathbf{N}_{\dot{p}}(\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\mathbf{L}\mathbf{c}_k=\mathbf{L}\mathbf{c}_k\\
&\iff(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma})\mathbf{I}_{(\dot{p},\dot{p})}(\boldsymbol{\Lambda}\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\mathbf{L}\mathbf{c}_k=\mathbf{L}\mathbf{c}_k\\
&\iff(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma})(\boldsymbol{\Gamma}'\otimes\boldsymbol{\Lambda}\boldsymbol{\Gamma}')\mathbf{L}\mathbf{c}_k=\mathbf{L}\mathbf{c}_k\ \textit{because}\ \mathbf{I}_{(\dot{p},\dot{p})}\mathbf{L}=\mathbf{L}\\
&\iff(\boldsymbol{\Psi}^{-1}\otimes\boldsymbol{\Psi})\mathbf{L}\mathbf{c}_k=\mathbf{L}\mathbf{c}_k\iff(\mathbf{I}_{\dot{p}}\otimes\boldsymbol{\Psi})\mathbf{L}\mathbf{c}_k=(\boldsymbol{\Psi}\otimes\mathbf{I}_{\dot{p}})\mathbf{L}\mathbf{c}_k\\
&\iff\boldsymbol{\Psi}\mathbf{D}_k=\mathbf{D}_k\boldsymbol{\Psi},
\end{aligned}$$

where $\mathbf{D}_k=\mathrm{Diag}(\mathbf{c}_k)$. *The above property (*$\boldsymbol{\Psi}$ *and* $\mathbf{D}_k$ *commute) along with symmetry of both matrices implies that* $\boldsymbol{\Psi}$ *and* $\mathbf{D}_j$ *can be diagonalized by the same orthogonal matrix (Schott, 1997, Theorem 4.15). To fulfill the condition that* $\mathrm{rank}\begin{pmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_k\end{pmatrix}=k$, *the vector* $\mathbf{c}_k$ *must assign two or more different values to coefficients that correspond to at least one of the* k *blocks. Without loss of generality, assume that the coefficient values associated with block* i *are not all identical. Choose* $\mathbf{Q}\in\mathcal{P}_{\dot{p}}$ *so that the diagonal elements of* $[\mathbf{Q}'\mathbf{D}_k\mathbf{Q}]_{ii}$ *are those associated with the* i^{th} *block and are sorted from smallest to largest. It follows that* $[\mathbf{Q}'\boldsymbol{\Psi}\mathbf{Q}]_{ii}=\mathbf{Q}_i'\boldsymbol{\Gamma}_{ii}^*\boldsymbol{\Lambda}_{ii}^*\boldsymbol{\Gamma}_{ii}^{*\prime}\mathbf{Q}_i$ *for some* $\mathbf{Q}_i\in\mathcal{P}_{p_i}$. *Denote the vector of eigenvalue multiplicities of* $[\mathbf{Q}'\mathbf{D}_k\mathbf{Q}]_{ii}$ *by* $\mathbf{m}^*=\begin{pmatrix}m_1^* & \cdots & m_{k^*}^*\end{pmatrix}'$, *where* $\mathbf{1}_{k^*}'\mathbf{m}^*=p_i$ *and* $k^*\geq 2$. *The eigenvectors of* $[\mathbf{Q}'\mathbf{D}_k\mathbf{Q}]_{ii}$ *that are associated with the* j *smallest eigenvalue can be written as* $\mathbf{E}_{j,\mathbf{m}^*}\mathbf{M}_j$, *where* $\mathbf{M}_j$ *is an arbitrary matrix chosen from* $\mathcal{O}_{m_j^*}$. *Because* $\boldsymbol{\Psi}$ *and* $\mathbf{D}_k$ *commute, the columns of* $\mathbf{E}_{j,\mathbf{m}^*}\mathbf{M}_j$ *also must be eigenvectors of*

$[\mathbf{Q}'\boldsymbol{\Psi}\mathbf{Q}]_{ii}$ *for some* $\mathbf{M}_j \in \mathcal{O}_{m_j^*}$. *Accordingly,*

$$\mathbf{Q}_i'\boldsymbol{\Gamma}_{ii}^*\boldsymbol{\Lambda}_{ii}^*\boldsymbol{\Gamma}_{ii}^{*\prime}\mathbf{Q}_i = \sum_{j=1}^{k^*}\mathbf{E}_{j,\mathbf{m}^*}\mathbf{M}_j\mathbf{E}_{j,\mathbf{m}^*}'\boldsymbol{\Lambda}_{ii}^*\mathbf{E}_{j,\mathbf{m}^*}\mathbf{M}_j'\mathbf{E}_{j,\mathbf{m}^*}' = \bigoplus_{j=1}^{k^*}\mathbf{M}_j\mathbf{E}_{j,\mathbf{m}^*}'\boldsymbol{\Lambda}_{ii}^*\mathbf{E}_{j,\mathbf{m}^*}\mathbf{M}_j'.$$

This contradicts the assumption that the maximal number of non-empty blocks is k *and, therefore, the rank of* $\mathbf{W}_{\psi}$ *cannot be less than* $k - \dot{p}$. □

In the remainder of this supplement, it is assumed that the $\dot{p}$ variables are ordered so that if $\mathbf{Q} = \mathbf{I}_{\dot{p}}$, then $\text{blk}(\mathbf{Q}) = k$ and, therefore, $\boldsymbol{\Psi}(\mathbf{Q})$ can be written more simply as $\boldsymbol{\Psi}$ and

$$\boldsymbol{\Psi} = \bigoplus_{i=1}^{k}\boldsymbol{\Psi}_i,$$

where $\boldsymbol{\Psi}_i$ is a $p_i \times p_i$ correlation matrix.

3 Proof of Theorem 2

Theorem 2. *Assume that* $\mathbf{W}_{\psi}$ *in* (12) *has full row-rank and write* $\mathbf{W}_{\psi}$ *in terms of its singular values and vectors:*

$$\mathbf{W}_{\psi} = \mathbf{U}_{\psi}\begin{pmatrix}\mathbf{D}_{\psi} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{A}_{\psi,1}' \\ \mathbf{A}_{\psi,2}'\end{pmatrix} = \mathbf{U}_{\psi}\,\mathbf{D}_{\psi}\,\mathbf{A}_{\psi,1}';$$

where $\mathbf{D}_{\psi}$ *is a* $(\dot{p}-1)\times(\dot{p}-1)$ *diagonal matrix of singular values;* $\mathbf{U}_{\psi} \in \mathcal{O}_{\dot{p}-1}$; $\mathbf{A}_{\psi,1}$ *has dimension* $(\dot{p}^2 - \mathbf{m}'\mathbf{m})/2 \times \dot{p} - 1$; $\mathbf{A}_{\psi,2}$ *has dimension* $(\dot{p}^2 - \mathbf{m}'\mathbf{m})/2 \times (\dot{p}^2 - \mathbf{m}'\mathbf{m})/2 - (\dot{p}-1)$; *and* $\mathbf{A}_{\psi} = \begin{pmatrix}\mathbf{A}_{\psi,1} & \mathbf{A}_{\psi,2}\end{pmatrix} \in \mathcal{O}_{(\dot{p}^2-\mathbf{m}'\mathbf{m})/2}$. *Let* $\mathbf{V}_1$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\psi,1})$, *the vector space generated by the columns of* $\mathbf{A}_{\psi,1}$. *Similarly, let* $\mathbf{V}_2$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\psi,2})$. *For fixed* $\mathbf{V} = \begin{pmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{pmatrix}$, *the matrices of first derivatives of* $\mathbf{g} = \text{vec}\,\mathbf{G}$ *with respect to* $\boldsymbol{\mu}'$ *and* $\boldsymbol{\varphi}'$ *are*

$$\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)} \overset{\text{def}}{=} \left.\frac{\partial\,\mathbf{g}}{\partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = (\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})})\mathbf{A}_2\mathbf{V}_2 \text{ and}$$

$$\mathbf{D}_{\mathbf{g}:\boldsymbol{\varphi}}^{(1)} \overset{\text{def}}{=} \left.\frac{\partial\,\mathbf{g}}{\partial\,\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = -\frac{1}{2}(\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})})\mathbf{A}_2\,\mathbf{W}_{\psi}^+\,\mathbf{C}_3'(\boldsymbol{\Gamma}\otimes\boldsymbol{\Gamma})\mathbf{L}\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)},$$

where $\mathbf{W}_{\psi}^+ = \mathbf{A}_{\psi,1}\,\mathbf{D}_{\psi}^{-1}\,\mathbf{U}_{\psi}' = \mathbf{W}_{\psi}'\left(\mathbf{W}_{\psi}\,\mathbf{W}_{\psi}'\right)^{-1}$ *is the Moore-Penrose inverse of* $\mathbf{W}_{\psi}$ *and* $\mathbf{L}$ *is given in (2).*

Proof. *For fixed* $\mathbf{V}$, *the derivative of* $\mathbf{g}$ *in (10) with respect to* $\boldsymbol{\mu}'$ *is*

$$\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)} = \mathbf{A}_1\left.\frac{\partial\,\boldsymbol{\eta}_{\boldsymbol{\mu},\boldsymbol{\varphi},1}}{\partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} + \mathbf{A}_2\mathbf{V}_1\left.\frac{\partial\,\boldsymbol{\eta}_{\boldsymbol{\mu},\boldsymbol{\varphi},2}}{\partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} + \mathbf{A}_2\mathbf{V}_2.$$

The derivatives of $\boldsymbol{\eta}_{\boldsymbol{\mu},\boldsymbol{\varphi}}$ *can be obtained by employing the constraints* $\mathbf{G}\mathbf{G}' = \mathbf{I}_{\dot{p}}$ *and (9). Differentiating the first constraint yields*

$$\left.\frac{\partial\,\text{vec}\,\mathbf{G}\mathbf{G}'}{\partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \mathbf{0} \Longrightarrow \mathbf{D}_{\dot{p}}'\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)} = \mathbf{0}$$

$$\Longrightarrow \mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)} = (\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})})\mathbf{A}_2\mathbf{V}_1\left.\frac{\partial\,\boldsymbol{\eta}_{\boldsymbol{\mu},\boldsymbol{\varphi},2}}{\partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} + (\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})})\mathbf{A}_2\mathbf{V}_2$$

because $\mathbf{A}_1$ *and* $\mathbf{A}_2$ *satisfy* $\mathbf{D}_{\dot{p}}'\mathbf{A}_1 = \mathbf{I}_{\dot{p}(\dot{p}+1)/2}$ *and* $\mathbf{A}_1\mathbf{D}_{\dot{p}}'\mathbf{A}_2 = \mathbf{I}_{(\dot{p},\dot{p})}\mathbf{A}_2$*. Differentiating the second constraint yields*

$$\left.\frac{\partial\, \mathbf{C}_3' \operatorname{vec} \mathbf{P}}{\partial\, \boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \mathbf{0} \Longrightarrow \mathbf{C}_3'(\boldsymbol{\Gamma}\boldsymbol{\Lambda} \otimes \boldsymbol{\Gamma})\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)} = \mathbf{0} \Longrightarrow$$

$$\left.\frac{\partial\, \boldsymbol{\eta}_{\boldsymbol{\mu},\boldsymbol{\varphi},2}}{\partial\, \boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = -\left(\mathbf{W}_{\boldsymbol{\psi}}\, \mathbf{V}_1\right)^{-1} \mathbf{W}_{\boldsymbol{\psi}}\, \mathbf{V}_2 = \mathbf{0}$$

because $\mathbf{A}_{\boldsymbol{\psi},1}'\, \mathbf{V}_2 = \mathbf{0}$*. Accordingly, the first derivative simplifies to* $(\mathbf{I}_{\dot{p}^2} - \mathbf{I}_{(\dot{p},\dot{p})})\mathbf{A}_2\mathbf{V}_2$ *as claimed. The proof for* $\mathbf{D}_{\mathbf{g}:\boldsymbol{\varphi}}^{(1)}$ *is similar to that of* $\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)}$ *and is omitted.* □

4 Proof of Theorem 3

Theorem 3. *Assume that* $\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi}}$ *has full row-rank and write* $\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi}}$ *in terms of its singular values and vectors:*

$$\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi}} = \mathbf{U}_{\boldsymbol{\xi}} \begin{pmatrix}\mathbf{D}_{\boldsymbol{\xi}} & \mathbf{0}\end{pmatrix} \begin{pmatrix}\mathbf{A}_{\boldsymbol{\xi},1}' \\ \mathbf{A}_{\boldsymbol{\xi},2}'\end{pmatrix} = \mathbf{U}_{\boldsymbol{\xi}}\mathbf{D}_{\boldsymbol{\xi}}\mathbf{A}_{\boldsymbol{\xi},1}';$$

where $\mathbf{W}_{\boldsymbol{\xi}}$ *is given in* (17)*;* $\mathbf{D}_{\boldsymbol{\xi}}$ *is an* $r_2 \times r_2$ *diagonal matrix of nonzero singular values;* $r_2 = \operatorname{rank}(\mathbf{C}_2)$*;* $\mathbf{U}_{\boldsymbol{\xi}} \in \mathcal{O}_{r_2}$*;* $\mathbf{A}_{\boldsymbol{\xi},1}$ *has dimension* $q_4 \times r_2$*;* $\mathbf{A}_{\boldsymbol{\xi},2}$ *has dimension* $q_4 \times \nu_3$*;* $\nu_3 = q_4 - r_2$*; and* $\begin{pmatrix}\mathbf{A}_{\boldsymbol{\xi},1} & \mathbf{A}_{\boldsymbol{\xi},2}\end{pmatrix} \in \mathcal{O}_{q_4}$*. Let* $\mathbf{V}_3$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\boldsymbol{\xi},1})$ *and let* $\mathbf{V}_4$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\boldsymbol{\xi},2})$*. For fixed* $\mathbf{V}^* = \begin{pmatrix}\mathbf{V}_3 & \mathbf{V}_4\end{pmatrix}$*, the first derivative of* $\boldsymbol{\lambda}$ *with respect to* $\boldsymbol{\varphi}'$ *is*

$$\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)} \overset{\text{def}}{=} \frac{\partial\, \boldsymbol{\lambda}}{\partial\, \boldsymbol{\varphi}'} = w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi}}\mathbf{V}_4, \ \ \textit{where } \mathbf{H}_{\boldsymbol{\lambda}} = \mathbf{I}_{\dot{p}} - \frac{1}{\dot{p}}\boldsymbol{\lambda}\mathbf{1}_{\dot{p}}'$$

and w *is defined in* (13)*.*

Proof. *For fixed* $\mathbf{V}_3$ *and* $\mathbf{V}_4$*, the first derivative of* $\boldsymbol{\lambda}$ *with respect to* $\boldsymbol{\varphi}'$ *can be written as*

$$\begin{aligned} \mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)} &= \left.\frac{\partial}{\partial \boldsymbol{\epsilon}'} \frac{\dot{p}\mathbf{T}_3 \exp\{\odot\mathbf{T}_4 \left[\mathbf{V}_3\boldsymbol{\eta}_{\boldsymbol{\varphi}} + \mathbf{V}_4(\boldsymbol{\varphi} + \boldsymbol{\epsilon})\right]\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3 \exp\{\odot\mathbf{T}_4 \left[\mathbf{V}_3\boldsymbol{\eta}_{\boldsymbol{\varphi}} + \mathbf{V}_4(\boldsymbol{\varphi} + \boldsymbol{\epsilon})\right]\}}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} \\ &= w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{W}_{\boldsymbol{\xi}} \left(\mathbf{V}_3 \left.\frac{\partial\, \boldsymbol{\eta}_{\boldsymbol{\varphi}}}{\partial\, \boldsymbol{\epsilon}'}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} + \mathbf{V}_4\right). \end{aligned}$$

An expression for $\partial\, \boldsymbol{\eta}_{\boldsymbol{\varphi}}/\partial\, \boldsymbol{\epsilon}'$ *can be obtained by setting the derivative of the constraint in* (16) *with respect to* $\boldsymbol{\epsilon}$ *to zero and solving for* $\partial\, \boldsymbol{\eta}_{\boldsymbol{\varphi}}/\partial\, \boldsymbol{\epsilon}'$*. The result is*

$$\left.\frac{\partial}{\partial\, \boldsymbol{\epsilon}'}\, \mathbf{C}_2' \exp\{\odot\mathbf{T}_4 \left[\mathbf{V}_3\boldsymbol{\eta}_{\boldsymbol{\varphi}} + \mathbf{V}_4(\boldsymbol{\varphi} + \boldsymbol{\epsilon})\right]\}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} = \mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi}} \left(\mathbf{V}_3 \left.\frac{\partial\, \boldsymbol{\eta}_{\boldsymbol{\varphi}}}{\partial\, \boldsymbol{\epsilon}'}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} + \mathbf{V}_4\right) = \mathbf{0}$$

$$\Longrightarrow \left.\frac{\partial\, \boldsymbol{\eta}_{\boldsymbol{\varphi}}}{\partial\, \boldsymbol{\epsilon}'}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} = -\left(\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi}}\mathbf{V}_3\right)^{-1} \mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi}}\mathbf{V}_4 = \mathbf{0}$$

because $\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1} = \mathbf{U}_{\boldsymbol{\xi}}\mathbf{D}_{\boldsymbol{\xi}}\mathbf{A}_{\boldsymbol{\xi},1}'$ *and* $\mathbf{A}_{\boldsymbol{\xi},1}'\mathbf{V}_4 = \mathbf{0}$*. Therefore,* $\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)}$ *simplifies to* $w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{W}_{\boldsymbol{\xi}}\mathbf{V}_4$ *as claimed.* □

5 Proof of Theorem 5

Theorem 5. *Define* $\mathbf{F}_0^{(1)}$ *and* $\mathbf{F}_{a,0}^{(1)}$ *as*

$$\mathbf{F}_0^{(1)} \stackrel{\text{def}}{=} \left.\frac{\partial \operatorname{vec} \boldsymbol{\Sigma}_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \text{ and } \mathbf{F}_{a,0}^{(1)} \stackrel{\text{def}}{=} \left.\frac{\partial \operatorname{vec} \boldsymbol{\Sigma}_a(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{a,0}}.$$

Denote the dimensions of Θ_0 *and* Θ_a *by* $\dot{\nu}_0$ *and* $\dot{\nu}_a$*, respectively. If*

(a) $\boldsymbol{\theta}_0$ *and* $\boldsymbol{\theta}_a$ *are in open subsets of their respective parameter spaces,*

(b) $\boldsymbol{\theta}_0$ *and* $\boldsymbol{\theta}_a$ *are consistent solutions to* $\partial Q/\partial\boldsymbol{\theta} = \mathbf{0}$*,*

(c) $\mathbf{F}_0^{(1)}$ *and* $\mathbf{F}_{a,0}^{(1)}$ *each have full column-rank, and*

(d) $\sqrt{n}(\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}) = O(1)$ *and* $\sqrt{n}(\boldsymbol{\Sigma}_a - \boldsymbol{\Sigma}) = O(1)$*, and*

(e) $\operatorname{vec} \mathbf{F}_0^{(1)} \in \mathcal{R}\left(\mathbf{I}_{\dot{\nu}_a} \otimes \mathbf{F}_{a,0}^{(1)}\right)$,

then

$$X^2 \sim \sum_{i=1}^{df} \alpha_i V_i + O_p\left(n^{-\frac{1}{2}}\right) \text{ for } Q_1 \text{ and } Q_2, \text{ and } X^2 \sim \chi^2_{df,\Delta} + O_p\left(n^{-\frac{1}{2}}\right) \text{ for } Q_3,$$

where $df = \dot{\nu}_a - \dot{\nu}_0$, V_i *for* $i = 1, \ldots, df$ *are independently distributed as* $V_i \sim \chi^2_{1,\Delta_i}$, α_i *is the* i^{th} *largest eigenvalue of*

$$\mathbf{T} = \frac{1}{2}(\boldsymbol{\Sigma}_0^{-1} \otimes \boldsymbol{\Sigma}_0^{-1})(\mathbf{P}_a - \mathbf{P}_0)\boldsymbol{\Omega}_n,$$

$$\mathbf{P}_0 = \mathbf{F}_0^{(1)}\left[\mathbf{F}_0^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1} \otimes \boldsymbol{\Sigma}_0^{-1}\right)\mathbf{F}_0^{(1)}\right]^{-1}\mathbf{F}_0^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1} \otimes \boldsymbol{\Sigma}_0^{-1}\right),$$

$$\mathbf{P}_a = \mathbf{F}_{a,0}^{(1)}\left[\mathbf{F}_{a,0}^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1} \otimes \boldsymbol{\Sigma}_0^{-1}\right)\mathbf{F}_{a,0}^{(1)}\right]^{-1}\mathbf{F}_{a,0}^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1} \otimes \boldsymbol{\Sigma}_0^{-1}\right), \quad \Delta_i = \frac{n}{2}\frac{\left[\mathbf{u}_i'(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\right]^2}{\mathbf{u}_i'\boldsymbol{\Omega}_n\mathbf{u}_i},$$

$\mathbf{u}_i$ *is an eigenvector of* $\mathbf{T}$ *that corresponds to* α_i, $\boldsymbol{\sigma}_0 = \operatorname{vec}\boldsymbol{\Sigma}_0$,

$$\Delta = \frac{n}{2}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)'\boldsymbol{\Omega}_n^+\left(\mathbf{P}_a^* - \mathbf{P}_0^*\right)(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0), \quad \mathbf{P}_0^* = \mathbf{F}_0^{(1)}\left[\mathbf{F}_0^{(1)\prime}\boldsymbol{\Omega}_n^+\mathbf{F}_0^{(1)}\right]^{-1}\mathbf{F}_0^{(1)\prime}\boldsymbol{\Omega}_n^+,$$

$$\mathbf{P}_a^* = \mathbf{F}_{a,0}^{(1)}\left[\mathbf{F}_{a,0}^{(1)\prime}\boldsymbol{\Omega}_n^+\mathbf{F}_{a,0}^{(1)}\right]^{-1}\mathbf{F}_{a,0}^{(1)\prime}\boldsymbol{\Omega}_n^+, \text{ and } \boldsymbol{\Omega}_n^+ = \mathbf{D}_{\dot{p}}\left(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}}\right)^{-1}\mathbf{D}_{\dot{p}}'.$$

If the multiplicity of α_i *is greater than one, then the corresponding eigenvectors are chosen to satisfy* $\mathbf{u}_i'\boldsymbol{\Omega}_n\mathbf{u}_j = 0$ *for all* $i \neq j$.

Proof. *The estimating function* $\partial Q(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$ *for the null model can be expanded around* $\widehat{\boldsymbol{\theta}}_0 = \boldsymbol{\theta}_0$ *as*

$$\frac{1}{\sqrt{n}}\left.\frac{\partial Q(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_0} = \mathbf{0} = \left(\frac{1}{\sqrt{n}}\left.\frac{\partial Q(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right) + \left(\frac{1}{n}\left.\frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \otimes \partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)\sqrt{n}(\widehat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}_0) + O_p\left(n^{-\frac{1}{2}}\right).$$

Accordingly,

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_0-\boldsymbol{\theta}_0)=-\left(\frac{1}{n}\frac{\partial^2 Q(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\otimes\partial\boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)^{-1}\left(\frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)+O_p\left(n^{-\frac{1}{2}}\right).$$

Make the substitution $\mathbf{S}^{-1}\otimes\mathbf{S}^{-1}=\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}+O_p\left(n^{-\frac{1}{2}}\right)$ *in* Q_2 *and* $\widehat{\boldsymbol{\Omega}}_n=\boldsymbol{\Omega}_n+O_p\left(n^{-\frac{1}{2}}\right)$ *in* Q_3*, and use* $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_0+O\left(n^{-\frac{1}{2}}\right)$ *to obtain*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_0-\boldsymbol{\theta}_0)=$$

$$\begin{cases}\left[\mathbf{F}_0^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)\mathbf{F}_0^{(1)}\right]^{-1}\mathbf{F}_0^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right) & \text{for } Q_1 \text{ and } Q_2,\\ \left[\mathbf{F}_0^{(1)\prime}\boldsymbol{\Omega}_n^{+}\mathbf{F}_0^{(1)}\right]^{-1}\mathbf{F}_0^{(1)\prime}\boldsymbol{\Omega}_n^{+}\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right) & \text{for } Q_3,\end{cases}$$

where $\boldsymbol{\Omega}_n^{+}=\mathbf{D}_{\dot{p}}(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{-1}\mathbf{D}_{\dot{p}}'$ *and* $\mathbf{D}_{\dot{p}}$ *is the duplication matrix. Expanding the estimating function* $\partial Q(\boldsymbol{\theta})/\partial\boldsymbol{\theta}=\mathbf{0}$ *for the alternative model around* $\widehat{\boldsymbol{\theta}}_a=\boldsymbol{\theta}_{a,0}$ *and solving for* $\sqrt{n}(\widehat{\boldsymbol{\theta}}_a-\boldsymbol{\theta}_{a,0})$ *yields*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_a-\boldsymbol{\theta}_{a,0})=$$

$$\begin{cases}\left[\mathbf{F}_{a,0}^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)\mathbf{F}_{a,0}^{(1)}\right]^{-1}\mathbf{F}_{a,0}^{(1)\prime}\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right) & \text{for } Q_1 \text{ and } Q_2,\\ \left[\mathbf{F}_{a,0}^{(1)\prime}\boldsymbol{\Omega}_n^{+}\mathbf{F}_{a,0}^{(1)}\right]^{-1}\mathbf{F}_{a,0}^{(1)\prime}\boldsymbol{\Omega}_n^{+}\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right) & \text{for } Q_3.\end{cases}$$

The second order Taylor expansion of $Q(\widehat{\boldsymbol{\theta}}_0)$ *around* $\widehat{\boldsymbol{\theta}}_0=\boldsymbol{\theta}_0$ *is*

$$Q(\widehat{\boldsymbol{\theta}}_0)=Q(\boldsymbol{\theta}_0)+\left(\frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)\sqrt{n}(\widehat{\boldsymbol{\theta}}_0-\boldsymbol{\theta}_0)$$

$$+\sqrt{n}(\widehat{\boldsymbol{\theta}}_0-\boldsymbol{\theta}_0)'\left(\frac{1}{n}\frac{\partial^2 Q(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\otimes\partial\boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right)\sqrt{n}(\widehat{\boldsymbol{\theta}}_0-\boldsymbol{\theta}_0)+O_p\left(n^{-\frac{1}{2}}\right).$$

Making the same substitutions as was done earlier and using the solution for $\sqrt{n}(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)$ *yields*

$$Q_1(\widehat{\boldsymbol{\theta}}_0)=n\left[\operatorname{tr}\left(\mathbf{S}\boldsymbol{\Sigma}_0^{-1}\right)+\ln|\boldsymbol{\Sigma}_0|\right]-\frac{1}{2}\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)'\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)\mathbf{P}_0\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right),$$

$$Q_2(\widehat{\boldsymbol{\theta}}_0)=\frac{1}{2}\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)'\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)\left(\mathbf{I}_{\dot{p}^2}-\mathbf{P}_0\right)\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right),\ \text{and}$$

$$Q_3(\widehat{\boldsymbol{\theta}}_0)=\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)'\boldsymbol{\Omega}_n^{+}\left(\mathbf{I}_{\dot{p}^2}-\mathbf{P}_0^{*}\right)\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right),$$

where $\mathbf{P}_0$ *and* $\mathbf{P}_0^{*}$ *are defined in Theorem 5. The expansions of* $Q_i(\widehat{\boldsymbol{\theta}}_a)$ *for* $i=1,2,3$ *under the alternative model have the same form as those under the null model, except that* $\mathbf{P}_a$ *is substituted for* $\mathbf{P}_0$ *and* $\mathbf{P}_a^{*}$ *is substituted for* $\mathbf{P}_a$*. The test statistic,* $X^2=Q(\widehat{\boldsymbol{\theta}}_0)-Q(\widehat{\boldsymbol{\theta}}_a)$*, therefore, can be written as*

$$X^2=\begin{cases}\frac{1}{2}\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)'\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)\left(\mathbf{P}_a-\mathbf{P}_0\right)\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right) & \text{for } Q_1 \text{ and } Q_2,\\ \sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)'\boldsymbol{\Omega}_n^{+}\left(\mathbf{P}_a^{*}-\mathbf{P}_0^{*}\right)\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0)+O_p\left(n^{-\frac{1}{2}}\right) & \text{for } Q_3.\end{cases}$$

Define $\mathbf{z}$ *as*

$$\mathbf{z} \overset{\text{def}}{=} (\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{-\frac{1}{2}}\mathbf{D}_{\dot{p}}'\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0).$$

Then,

$$\mathbf{z} + O_p\left(n^{-\frac{1}{2}}\right) \sim \mathrm{N}\left[(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{-\frac{1}{2}}\mathbf{D}_{\dot{p}}'\sqrt{n}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_0), \mathbf{I}_{\dot{p}(\dot{p}+1)/2}\right].$$

Furthermore, it follows from $\mathbf{N}_{\dot{p}}(\mathbf{s}-\boldsymbol{\sigma}_0) = (\mathbf{s}-\boldsymbol{\sigma}_0)$ *that*

$$\sqrt{n}(\mathbf{s}-\boldsymbol{\sigma}_0) = \mathbf{D}_{\dot{p}}^{+\prime}(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{\frac{1}{2}}\mathbf{z},$$

where $\mathbf{D}_{\dot{p}}^{+} = (\mathbf{D}_{\dot{p}}'\mathbf{D}_{\dot{p}})^{-1}\mathbf{D}_{\dot{p}}'$. *Accordingly, the test statistic can be written as*

$$X^2 =$$

$$\begin{cases} \frac{1}{2}\mathbf{z}'(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{\frac{1}{2}}\mathbf{D}_{\dot{p}}^{+}\left(\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Sigma}_0^{-1}\right)(\mathbf{P}_a-\mathbf{P}_0)\,\mathbf{D}_{\dot{p}}^{+\prime}(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{\frac{1}{2}}\mathbf{z} + O_p\left(n^{-\frac{1}{2}}\right) & \textit{for } Q_1 \textit{ and } Q_2, \\ \mathbf{z}'(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{\frac{1}{2}}\mathbf{D}_{\dot{p}}^{+}\boldsymbol{\Omega}_n^{+}\left(\mathbf{P}_a^{*}-\mathbf{P}_0^{*}\right)\mathbf{D}_{\dot{p}}^{+\prime}(\mathbf{D}_{\dot{p}}'\boldsymbol{\Omega}_n\mathbf{D}_{\dot{p}})^{\frac{1}{2}}\mathbf{z} + O_p\left(n^{-\frac{1}{2}}\right) & \textit{for } Q_3. \end{cases}$$

The proof is completed by using the asymptotic normal distribution of $\mathbf{z}$ *and standard results on quadratic forms of normal random vectors (e.g., Stapleton, 1995, p. 65).* □

6 Higher-Order Derivatives

In this section, second- and third-order derivatives of model components with respect to $\boldsymbol{\theta}$ are listed. Higher-order derivatives are constructed in the same manner as in Theorems 2 and 3. Proofs are not given. For completeness, the first-order derivatives also are listed.

6.1 Arrangement of Derivatives

Let $\mathbf{A} = \mathbf{A}(\mathbf{B})$ be an $a \times q$ matrix whose elements are functions of the $b \times s$ matrix $\mathbf{B}$. Then the derivative of $\mathbf{A}$ with respect to $\mathbf{B}$ is defined to be $ab \times qs$ matrix of derivatives $\{\partial\, a_{ij}/\partial\, b_{st}\}$ whose elements are arranged as follows:

$$\frac{\partial\mathbf{A}}{\partial\mathbf{B}} \overset{\text{def}}{=} \frac{\partial}{\partial\mathbf{B}}\otimes\mathbf{A}.$$

Partition $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = \begin{pmatrix}\boldsymbol{\theta}_1' & \cdots & \boldsymbol{\theta}_k'\end{pmatrix}'$. In particular, if $\boldsymbol{\theta}$ has the structure in (3) then $k = 3$ and $\boldsymbol{\theta} = \begin{pmatrix}\boldsymbol{\tau}' & \boldsymbol{\mu}' & \boldsymbol{\varphi}'\end{pmatrix}'$. If the extended parameterization (§5.2 of the article) is used, then $k = 4$ and $\boldsymbol{\theta} = \begin{pmatrix}\boldsymbol{\tau}' & \boldsymbol{\mu}' & \boldsymbol{\varphi}' & \boldsymbol{\beta}'\end{pmatrix}'$. Derivatives of $\operatorname{vec}\boldsymbol{\Sigma}$ with respect to $\boldsymbol{\theta}$ are arranged as follows:

$$\mathbf{F}^{(1)} \overset{\text{def}}{=} \left.\frac{\partial\operatorname{vec}(\boldsymbol{\Sigma})}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \sum_{s=1}^{k}\left.\frac{\partial\operatorname{vec}(\boldsymbol{\Sigma})}{\partial\boldsymbol{\theta}_s'}\mathbf{E}_{s,\boldsymbol{\nu}}'\right|_{\boldsymbol{\mu}=\mathbf{0}};$$

$$\mathbf{F}^{(2)} \overset{\text{def}}{=} \left.\frac{\partial^2\operatorname{vec}(\boldsymbol{\Sigma})}{\partial\boldsymbol{\theta}'\otimes\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \sum_{s=1}^{k}\sum_{t=1}^{k}\left.\frac{\partial^2\operatorname{vec}(\boldsymbol{\Sigma})}{\partial\boldsymbol{\theta}_s'\otimes\partial\boldsymbol{\theta}_t'}\left(\mathbf{E}_{s,\boldsymbol{\nu}}\otimes\mathbf{E}_{t,\boldsymbol{\nu}}\right)'\right|_{\boldsymbol{\mu}=\mathbf{0}};$$

$$\mathbf{F}^{(11)} \overset{\text{def}}{=} \left.\frac{\partial^2\operatorname{vec}(\boldsymbol{\Sigma})}{\partial\boldsymbol{\theta}\otimes\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \operatorname{dvec}\left(\mathbf{F}^{(2)}, \dot{p}^2\dot{\nu}, \dot{\nu}\right);$$

$$\mathbf{F}^{(3)} \overset{\text{def}}{=} \left.\frac{\partial^3\operatorname{vec}(\boldsymbol{\Sigma})}{\partial\boldsymbol{\theta}'\otimes\partial\boldsymbol{\theta}'\otimes\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\mu}=\mathbf{0}}$$

$$= \sum_{s=1}^{k}\sum_{t=1}^{k}\sum_{u=1}^{k} \frac{\partial^3 \operatorname{vec}(\mathbf{\Sigma})}{\partial\boldsymbol{\theta}'_s \otimes \partial\boldsymbol{\theta}'_t \otimes \partial\boldsymbol{\theta}'_u} \left(\mathbf{E}_{s,\boldsymbol{\nu}} \otimes \mathbf{E}_{t,\boldsymbol{\nu}} \otimes \mathbf{E}_{u,\boldsymbol{\nu}}\right)' \bigg|_{\boldsymbol{\mu}=\mathbf{0}} ; \text{ and}$$

$$\mathbf{F}^{(111)} \overset{\text{def}}{=} \frac{\partial^3 \operatorname{vec}(\mathbf{\Sigma})}{\partial\boldsymbol{\theta} \otimes \partial\boldsymbol{\theta} \otimes \partial\boldsymbol{\theta}'}\bigg|_{\boldsymbol{\mu}=\mathbf{0}} = \operatorname{dvec}\left(\mathbf{F}^{(3)}, \dot{p}^2\dot{\nu}^2, \dot{\nu}\right);$$

where $\mathbf{E}_{s,\boldsymbol{\nu}}$ is defined in (4);

$$\boldsymbol{\nu} = \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_k \end{pmatrix} = \begin{pmatrix} \dim(\boldsymbol{\theta}_1) \\ \vdots \\ \dim(\boldsymbol{\theta}_k) \end{pmatrix}; \quad \dot{\nu} = \sum_{j=1}^{k} \nu_j;$$

and $\operatorname{dvec}(\mathbf{A}, a, b)$ is the $a \times b$ matrix that satisfies $\operatorname{vec}(\mathbf{A}) = \operatorname{vec}\{\operatorname{dvec}(\mathbf{A}, a, b)\}$. The matrix $\mathbf{A}$ must have exactly ab elements. Higher-order derivatives can be rearranged by post-multiplying by commutation matrices. For example,

$$\begin{aligned}
\frac{\partial^2 \operatorname{vec}(\mathbf{\Sigma})}{\partial\boldsymbol{\theta}'_t \otimes \partial\boldsymbol{\theta}'_s} &= \frac{\partial^2 \operatorname{vec}(\mathbf{\Sigma})}{\partial\boldsymbol{\theta}'_s \otimes \partial\boldsymbol{\theta}'_t} \mathbf{I}_{(\nu_t,\nu_s)} \text{ and} \\
\frac{\partial^3 \operatorname{vec}(\mathbf{\Sigma})}{\partial\boldsymbol{\theta}'_t \otimes \partial\boldsymbol{\theta}'_s \otimes \partial\boldsymbol{\theta}'_u} &= \frac{\partial^3 \operatorname{vec}(\mathbf{\Sigma})}{\partial\boldsymbol{\theta}'_s \otimes \partial\boldsymbol{\theta}'_t \otimes \partial\boldsymbol{\theta}'_u} \left(\mathbf{I}_{(\nu_t,\nu_s)} \otimes \mathbf{I}_{\nu_u}\right) \\
&= \frac{\partial^3 \operatorname{vec}(\mathbf{\Sigma})}{\partial\boldsymbol{\theta}'_t \otimes \partial\boldsymbol{\theta}'_u \otimes \partial\boldsymbol{\theta}'_s} \left(\mathbf{I}_{\nu_t} \otimes \mathbf{I}_{(\nu_s,\nu_u)}\right).
\end{aligned}$$

The derivatives of $\operatorname{vec}\mathbf{\Sigma}$ with respect to $\boldsymbol{\theta}$ satisfy the following multiplicative invariance properties:

$$\begin{aligned}
\mathbf{F}^{(1)} &= \mathbf{I}_{(\dot{p},\dot{p})}\mathbf{F}^{(1)}; \\
\mathbf{F}^{(2)} &= \mathbf{I}_{(\dot{p},\dot{p})}\,\mathbf{F}^{(2)} = \mathbf{F}^{(2)}\,\mathbf{I}_{(\dot{\nu},\dot{\nu})}; \\
\mathbf{F}^{(11)} &= \left(\mathbf{I}_{\dot{\nu}} \otimes \mathbf{I}_{(\dot{p},\dot{p})}\right)\mathbf{F}^{(11)}; \\
\mathbf{F}^{(3)} &= \mathbf{I}_{(\dot{p},\dot{p})}\,\mathbf{F}^{(3)} = \mathbf{F}^{(3)}\,\mathbf{I}_{(\dot{\nu},\dot{\nu}^2)} = \mathbf{F}^{(3)}\,\mathbf{I}_{(\dot{\nu}^2,\dot{\nu})}; \\
&= \mathbf{F}^{(3)}(\mathbf{I}_{\dot{\nu}} \otimes \mathbf{I}_{(\dot{\nu},\dot{\nu})}) = \mathbf{F}^{(3)}(\mathbf{I}_{(\dot{\nu},\dot{\nu})} \otimes \mathbf{I}_{\dot{\nu}}); \text{ and} \\
\mathbf{F}^{(111)} &= \left(\mathbf{I}_{(\dot{p},\dot{p})} \otimes \mathbf{I}_{\dot{\nu}^2}\right)\mathbf{F}^{(111)} = \left(\mathbf{I}_{\dot{p}^2} \otimes \mathbf{I}_{(\dot{\nu},\dot{\nu})}\right)\mathbf{F}^{(111)}.
\end{aligned}$$

These properties are useful when simplifying expressions.

The derivatives of $\boldsymbol{\lambda}$, $\operatorname{vec}\mathbf{G}$, and $\operatorname{vec}\mathbf{\Sigma}$ involve several structured matrices. For ease of reference, these structured matrices are defined below:

$$\mathbf{N}_a = \frac{1}{2}\left(\mathbf{I}_{a^2} + \mathbf{I}_{(a,a)}\right); \qquad \mathbf{I}^{(2)}_{\dot{p},\boldsymbol{\lambda}} = \left(\mathbf{\Lambda} \otimes \mathbf{I}_{\dot{p}}\right);$$

$$\mathbf{N}^*_a = \frac{1}{2}\left(\mathbf{I}_{a^2} - \mathbf{I}_{(a,a)}\right); \qquad \mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i} = \left(\mathbf{\Lambda}_i^{-1} \otimes \mathbf{I}_{p_i}\right);$$

$$\mathbf{J}_a = \mathbf{I}_{a^3} + \mathbf{I}_{(a,a^2)} + (\mathbf{I}_a \otimes \mathbf{I}_{(a,a)}); \qquad \mathbf{I}^{(3)}_{\dot{p}} = \left(\mathbf{I}_{\dot{p}} \otimes \operatorname{vec}\mathbf{I}_{\dot{p}} \otimes \mathbf{I}_{\dot{p}}\right)';$$

$$
\begin{aligned}
\mathbf{J}^*_a &= \mathbf{I}_{a^3} + \mathbf{I}_{(a,a^2)} + (\mathbf{I}_{(a,a)} \otimes \mathbf{I}_a); & \mathbf{I}^{(3)}_{p_i} &= (\mathbf{I}_{p_i} \otimes \operatorname{vec} \mathbf{I}_{p_i} \otimes \mathbf{I}_{p_i})'\,; \qquad (34)\\
\mathbf{J}_{a,b} &= \mathbf{I}_{(b,ab)} + (\mathbf{I}_b \otimes \mathbf{I}_{(b,a)}); & \mathbf{I}^{(3)}_{\dot{p},\boldsymbol{\lambda}} &= (\mathbf{I}_{\dot{p}} \otimes \operatorname{vec} \boldsymbol{\Lambda} \otimes \mathbf{I}_{\dot{p}})'\\
\mathbf{J}^*_{a,b} &= \mathbf{I}_{ab^2} + (\mathbf{I}_{(b,b)} \otimes \mathbf{I}_a); & \mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i} &= (\mathbf{I}_{p_i} \otimes \operatorname{vec} \boldsymbol{\Lambda}_i \otimes \mathbf{I}_{p_i})'\\
\mathbf{J}^{**}_{a,b} &= \mathbf{I}_{a^2b} + (\mathbf{I}_b \otimes \mathbf{I}_{(a,a)}); & \mathbf{I}^{(4)}_{\dot{p}} &= (\mathbf{I}_{\dot{p}} \otimes \operatorname{vec} \mathbf{I}_{\dot{p}} \otimes \operatorname{vec} \mathbf{I}_{\dot{p}} \otimes \mathbf{I}_{\dot{p}})'\,;\\
& & \mathbf{I}^{(4)}_{p_i} &= (\mathbf{I}_{p_i} \otimes \operatorname{vec} \mathbf{I}_{p_i} \otimes \operatorname{vec} \mathbf{I}_{p_i} \otimes \mathbf{I}_{p_i})'\,;
\end{aligned}
$$

where a and b are positive integers, and $\mathbf{I}_{(\cdot,\cdot)}$ is the commutation matrix (MacRae, 1974). The commutation matrix $\mathbf{I}_{(a,b)}$ is denoted by some writers as $\mathbf{K}_{b,a}$ (Magnus & Neudecker 1979, 1999 §3.7). The matrix $\mathbf{N}_a$ is the perpendicular projection operator that projects onto the vector space generated by symmetric matrices. That is, $\mathbf{N}_a' = \mathbf{N}_a$, $\mathbf{N}_a^2 = \mathbf{N}_a$, and if $\mathbf{A}$ is an $a \times a$ matrix, then $\operatorname{vec} \mathbf{A} = \mathbf{N}_a \operatorname{vec} \mathbf{A}$ if and only if $\mathbf{A}$ is symmetric. The matrix $\mathbf{N}_a^*$ is the perpendicular projection operator that projects onto the vector space generated by skew-symmetric matrices. That is, $\mathbf{N}_a^{*\prime} = \mathbf{N}_a^*$, $\mathbf{N}_a^{*2} = \mathbf{N}_a^*$ and if $\mathbf{A}$ is an $a \times a$ matrix, then $\operatorname{vec} \mathbf{A} = \mathbf{N}_a^* \operatorname{vec} \mathbf{A}$ if and only if $\mathbf{A}$ is skew-symmetric.

6.2 Derivatives of Scale Parameters

The vector of scale parameters, $\boldsymbol{\sigma}_d$, is parameterized as

$$\boldsymbol{\sigma}_d = \mathbf{T}_1 \exp\{\odot \mathbf{T}_2 \boldsymbol{\tau}\},$$

where $\mathbf{T}_1$ and $\mathbf{T}_2$ are full column-rank design matrices with dimensions $\dot{p} \times q_1$ and $q_1 \times \nu_1$, respectively.

Theorem 6. *The first three derivatives of $\boldsymbol{\sigma}_d$ with respect to $\boldsymbol{\tau}'$ are*

$$
\begin{aligned}
\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}} &= \frac{\partial\, \boldsymbol{\sigma}_d}{\partial\, \boldsymbol{\tau}'} = \mathbf{T}_1 \operatorname{Diag}\left(\exp\{\odot \mathbf{T}_2 \boldsymbol{\tau}\}\right) \mathbf{T}_2,\\
\mathbf{D}^{(2)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau},\boldsymbol{\tau}} &= \frac{\partial^2 \boldsymbol{\sigma}_d}{\partial\, \boldsymbol{\tau}' \otimes \partial\, \boldsymbol{\tau}'} = \mathbf{T}_1 \sum_{i=1}^{q_1} \mathbf{e}_i^{q_1} \exp\{\mathbf{e}_i^{q_1\prime} \mathbf{T}_2 \boldsymbol{\tau}\} (\mathbf{e}_i^{q_1\prime} \mathbf{T}_2 \otimes \mathbf{e}_i^{q_1\prime} \mathbf{T}_2), \ \textit{and}\\
\mathbf{D}^{(3)}_{\boldsymbol{\sigma}:\boldsymbol{\tau},\boldsymbol{\tau},\boldsymbol{\tau}} &= \frac{\partial^3 \boldsymbol{\sigma}_d}{\partial\, \boldsymbol{\tau}' \otimes \partial\, \boldsymbol{\tau}' \otimes \partial\, \boldsymbol{\tau}'} = \mathbf{T}_1 \sum_{i=1}^{q_1} \mathbf{e}_i^{q_1} \exp\{\mathbf{e}_i^{q_1\prime} \mathbf{T}_2 \boldsymbol{\tau}\} (\mathbf{e}_i^{q_1\prime} \mathbf{T}_2 \otimes \mathbf{e}_i^{q_1\prime} \mathbf{T}_2 \otimes \mathbf{e}_i^{q_1\prime} \mathbf{T}_2),
\end{aligned}
$$

where $\mathbf{e}_i^{q_1}$ is the i^{th} column of $\mathbf{I}_{q_1}$; □

6.3 Derivatives of Eigenvectors of $\boldsymbol{\Psi}$

The correlation matrix $\boldsymbol{\Psi}$ can be written as

$$\boldsymbol{\Psi} \;=\; \boldsymbol{\Gamma} \mathbf{G} \boldsymbol{\Lambda} \mathbf{G}' \boldsymbol{\Gamma}' \Big|_{\boldsymbol{\mu}=\mathbf{0}} = \bigoplus_{i=1}^{k} \boldsymbol{\Psi}_i, \ \text{where} \qquad (35)$$

$$\begin{aligned}\boldsymbol{\Psi}_i &= \boldsymbol{\Gamma}_i\mathbf{G}_i\boldsymbol{\Lambda}_i\mathbf{G}_i'\boldsymbol{\Gamma}_i\Big|_{\boldsymbol{\mu}_i=\mathbf{0}},\\ \boldsymbol{\Gamma} &= \bigoplus_{i=1}^{k}\boldsymbol{\Gamma}_i, \quad \mathbf{G}=\bigoplus_{i=1}^{k}\mathbf{G}_i, \text{ and } \mathbf{G}_i=\mathbf{G}_i(\boldsymbol{\mu}_i).\end{aligned}$$

Denote the vector of eigenvalue multiplicities that corresponds to $\boldsymbol{\Psi}_i$ by $\mathbf{m}_i$ and let $\mathbf{V}_i$ be a $(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2\times(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2$ nonsingular matrix that can be partitioned as $\mathbf{V}_i=\begin{pmatrix}\mathbf{V}_{1i} & \mathbf{V}_{2i}\end{pmatrix}$, where $\mathbf{V}_{1i}$ has dimension $(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2\times(p_i-1)$, $\mathbf{V}_{2i}$ has dimension $(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2\times\nu_{2i}$, $\nu_{2i}=[(p_i-1)^2+1-\mathbf{m}_i'\mathbf{m}_i]/2$, and $\mathbf{V}_{1i}'\mathbf{V}_{2i}=\mathbf{0}$. Then, $\operatorname{vec}\mathbf{G}_i$ can be written as

$$\mathbf{g}_i \overset{\text{def}}{=} \operatorname{vec}\mathbf{G}_i=\mathbf{A}_{1i}\boldsymbol{\eta}_i^*+\mathbf{A}_{2i}\boldsymbol{\mu}_i^*=\begin{pmatrix}\mathbf{A}_{1i} & \mathbf{A}_{2i}\mathbf{V}_{1i}\end{pmatrix}\begin{pmatrix}\boldsymbol{\eta}_{1i}\\ \boldsymbol{\eta}_{2i}\end{pmatrix}+\mathbf{A}_{2i}\mathbf{V}_{2i}\boldsymbol{\mu}_i, \tag{36}$$

where $\boldsymbol{\eta}_{1i}=\boldsymbol{\eta}_i^*$; $\boldsymbol{\eta}_{2i}=(\mathbf{V}_{1i}'\mathbf{V}_{1i})^{-1}\mathbf{V}_{1i}'\boldsymbol{\mu}_i^*$; $\boldsymbol{\mu}_i=(\mathbf{V}_{2i}'\mathbf{V}_{2i})^{-1}\mathbf{V}_{2i}'\boldsymbol{\mu}_i^*$; and $\mathbf{A}_{1i}$ and $\mathbf{A}_{2i}$ are known semi-orthogonal indicator matrices of order $p_i^2\times p_i(p_i+1)/2$ and $p_i^2\times(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2$, respectively. These matrices satisfy $\mathbf{A}_{1i}'\operatorname{vec}\mathbf{G}_i=\boldsymbol{\eta}_i^*$, $\mathbf{A}_{2i}'\operatorname{vec}\mathbf{G}_i=\boldsymbol{\mu}_i^*$, and $\mathbf{A}_{1i}'\mathbf{A}_{2i}=\mathbf{0}$ (see §2.4 of the article).

Expressions for the first derivatives of $\mathbf{g}=\operatorname{vec}\mathbf{G}$ with respect to $\boldsymbol{\mu}'$ and $\boldsymbol{\varphi}'$ for specific choices of $\mathbf{V}_i$ for $i=1,\ldots,k$ are given in Theorem 7. The following definitions are used in Theorem 7 and throughout this supplement:

$$\mathbf{L}_a=\sum_{i=1}^{a}(\mathbf{e}_i^a\otimes\mathbf{e}_i^a)\mathbf{e}_i^{a\prime} \text{ and } \mathbf{C}_{3i}=\sum_{j=1}^{p_i-1}(\mathbf{e}_j^{p_i}\otimes\mathbf{e}_j^{p_i})\mathbf{e}_j^{p_i\prime}, \tag{37}$$

where a is an integer.

Theorem 7. *Define* $\mathbf{W}_{\boldsymbol{\psi},i}$ *as* $\mathbf{W}_{\boldsymbol{\psi},i}=2\mathbf{C}_{3i}'(\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\otimes\boldsymbol{\Gamma}_i)\mathbf{N}_{p_i}^*\mathbf{A}_{2i}$*, where* $\mathbf{N}_{p_i}^*$ *is given in (34). Express* $\mathbf{W}_{\boldsymbol{\psi},i}$ *in terms of its singular values and vectors:*

$$\mathbf{W}_{\boldsymbol{\psi},i}=\mathbf{U}_{\boldsymbol{\psi},i}\begin{pmatrix}\mathbf{D}_{\boldsymbol{\psi},i} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{A}_{\boldsymbol{\psi},1i}'\\ \mathbf{A}_{\boldsymbol{\psi},2i}'\end{pmatrix}=\mathbf{U}_{\boldsymbol{\psi},i}\,\mathbf{D}_{\boldsymbol{\psi},i}\,\mathbf{A}_{\boldsymbol{\psi},1i}';$$

where $\mathbf{D}_{\boldsymbol{\psi},i}$ *is a* $(p_i-1)\times(p_i-1)$ *diagonal matrix of singular values;* $\mathbf{U}_{\boldsymbol{\psi},i}\in\mathcal{O}(p_i-1)$*;* $\mathbf{A}_{\boldsymbol{\psi},1i}$ *has dimension* $(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2\times p_i-1$*;* $\mathbf{A}_{\boldsymbol{\psi},2i}$ *has dimension* $(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2\times(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2-p_i+1$*; and* $\mathbf{A}_{\boldsymbol{\psi},i}=\begin{pmatrix}\mathbf{A}_{\boldsymbol{\psi},1i} & \mathbf{A}_{\boldsymbol{\psi},2i}\end{pmatrix}\in\mathcal{O}[(p_i^2-\mathbf{m}_i'\mathbf{m}_i)/2]$*. Let* $\mathbf{V}_{1i}$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\boldsymbol{\psi},1i})$ *and let* $\mathbf{V}_{2i}$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\boldsymbol{\psi},2i})$*. This choice of* $\mathbf{V}_i=\begin{pmatrix}\mathbf{V}_{1i} & \mathbf{V}_{2i}\end{pmatrix}$ *is delicate because* $\mathbf{V}_i$ *is a function of* $\boldsymbol{\rho}_i=\operatorname{vec}\boldsymbol{\Psi}_i$*, yet it will be treated as a constant when taking derivatives with respect to* $\boldsymbol{\varphi}$*. For fixed* $\mathbf{V}_i$*, the first three derivatives of* $\mathbf{g}=\operatorname{vec}\mathbf{G}$ *with respect to* $\boldsymbol{\mu}'$ *and* $\boldsymbol{\varphi}'$ *are*

$$\begin{aligned}\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)} &= \frac{\partial\operatorname{vec}\mathbf{G}}{\partial\boldsymbol{\mu}'}\bigg|_{\boldsymbol{\mu}=\mathbf{0}}=\sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}}\otimes\mathbf{E}_{i,\mathbf{p}})\mathbf{D}_{\mathbf{g}_i:\boldsymbol{\mu}_i}^{(1)}\mathbf{E}_{i,\boldsymbol{\nu}_2}';\\ \boldsymbol{\nu}_2 &= \begin{pmatrix}\nu_{21} & \cdots & \nu_{2k}\end{pmatrix}';\nu_{2i}=\frac{1}{2}\left[(p_i-1)^2+1-\mathbf{m}_i'\mathbf{m}_i\right];\\ \mathbf{D}_{\mathbf{g}_i:\boldsymbol{\mu}_i}^{(1)} &= \frac{\partial\operatorname{vec}\mathbf{G}_i}{\partial\boldsymbol{\mu}_i'}\bigg|_{\boldsymbol{\mu}_i=\mathbf{0}}=2\mathbf{N}_{p_i}^*\mathbf{A}_{2i}\mathbf{V}_{2i};\\ \mathbf{D}_{\mathbf{g}:\boldsymbol{\mu},\boldsymbol{\mu}}^{(2)} &= \frac{\partial^2\operatorname{vec}\mathbf{G}}{\partial\boldsymbol{\mu}'\otimes\partial\boldsymbol{\mu}'}\bigg|_{\boldsymbol{\mu}=\mathbf{0}}=\sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}}\otimes\mathbf{E}_{i,\mathbf{p}})\mathbf{D}_{\mathbf{g}_i:\boldsymbol{\mu}_i,\boldsymbol{\mu}_i}^{(2)}(\mathbf{E}_{i,\boldsymbol{\nu}_2}\otimes\mathbf{E}_{i,\boldsymbol{\nu}_2})';\end{aligned}$$

$$\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\mu}_i,\boldsymbol{\mu}_i} = \left.\frac{\partial^2 \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\mu}_i'\otimes\partial\,\boldsymbol{\mu}_i'}\right|_{\boldsymbol{\mu}_i=\mathbf{0}} = \Big[(\mathbf{I}-\mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}'_{p_i}\mathbf{I}^{(3)}_{p_i} + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i}\Big]\Big(\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\otimes\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\Big);$$

$$\mathbf{H}_{\boldsymbol{\psi},i} = 2\mathbf{N}^*_{p_i}\mathbf{A}_{2i}\,\mathbf{W}^+_{\boldsymbol{\psi},i}\,\mathbf{C}'_{3i}(\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\otimes\boldsymbol{\Gamma}_i);$$

$$\mathbf{W}^+_{\boldsymbol{\psi},i} = \mathbf{A}_{\boldsymbol{\psi},3i}\,\mathbf{D}^{-1}_{\boldsymbol{\psi},i}\,\mathbf{U}^{'}_{\boldsymbol{\psi},i}\ \ \textit{is the Moore-Penrose inverse of}\ \mathbf{W}_{\boldsymbol{\psi},i};$$

$$\begin{aligned}
\mathbf{D}^{(3)}_{\mathbf{g}:\boldsymbol{\mu},\boldsymbol{\mu},\boldsymbol{\mu}} &= \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}}{\partial\,\boldsymbol{\mu}'\otimes\partial\,\boldsymbol{\mu}'\otimes\partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}}\\
&= \sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}}\otimes\mathbf{E}_{i,\mathbf{p}})\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\mu}_i,\boldsymbol{\mu}_i,\boldsymbol{\mu}_i}(\mathbf{E}_{i,\boldsymbol{\nu}_2}\otimes\mathbf{E}_{i,\boldsymbol{\nu}_2}\otimes\mathbf{E}_{i,\boldsymbol{\nu}_2})';
\end{aligned}$$

$$\begin{aligned}
\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\mu}_i,\boldsymbol{\mu}_i,\boldsymbol{\mu}_i} &= \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\mu}_i'\otimes\partial\,\boldsymbol{\mu}_i'\otimes\partial\,\boldsymbol{\mu}_i'}\right|_{\boldsymbol{\mu}_i=\mathbf{0}}\\
&= -\Big[(\mathbf{I}-\mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}'_{p_i}\mathbf{I}^{(3)}_{p_i} + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i}\Big]\Big(\mathbf{I}_{(p_i,p_i)}\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\mu}_i,\boldsymbol{\mu}_i}\otimes\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\Big)\mathbf{J}_{\nu_{2i}};
\end{aligned}$$

$$\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}} = \left.\frac{\partial\ \operatorname{vec}\mathbf{G}}{\partial\,\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}}\otimes\mathbf{E}_{i,\mathbf{p}})\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}};$$

$$\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}} = \left.\frac{\partial\ \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}_i=\mathbf{0}} = -\frac{1}{2}\,\mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}};$$

$$\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi}} = \left.\frac{\partial^2 \operatorname{vec}\mathbf{G}}{\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}}\otimes\mathbf{E}_{i,\mathbf{p}})\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi}};$$

$$\begin{aligned}
\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi}} &= \left.\frac{\partial^2 \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}_i=\mathbf{0}} = \Big[(\mathbf{I}-\mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}'_{p_i}\mathbf{I}^{(3)}_{p_i} + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i}\Big]\Big(\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}}\Big)\\
&\quad +2\,\mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i}\Big(\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}}\otimes\mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\Big)\mathbf{N}_{\nu_3} - \frac{1}{2}\,\mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}};
\end{aligned}$$

$$\mathbf{D}^{(3)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}} = \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}}{\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}}\otimes\mathbf{E}_{i,\mathbf{p}})\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}};$$

$$\begin{aligned}
\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}} &= \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}_i=\mathbf{0}}\\
&= -\Big[(\mathbf{I}-\mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}'_{p_i}\mathbf{I}^{(3)}_{p_i} + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i}\Big]\Big(\mathbf{I}_{(p_i,p_i)}\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}}\Big)\mathbf{J}_{\nu_3}
\end{aligned}$$

$$
\begin{aligned}
& + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(4)}_{p_i}\left(\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}} \otimes \mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}}\right)\mathbf{J}^{*}_{\nu_3} - \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i} \\
& \times \Big[\left(\mathbf{I}_{(p_i,p_i)}\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi}} \otimes \mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\right) + \left(\mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}}\right)\Big]\mathbf{J}_{\nu_3} \\
& - \frac{1}{2}\,\mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(3)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}};
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\mu}} &= \left.\frac{\partial^2 \operatorname{vec}\mathbf{G}}{\partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}} \otimes \mathbf{E}_{i,\mathbf{p}})\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\mu}_i}(\mathbf{I}_{\nu_3} \otimes \mathbf{E}'_{i,\boldsymbol{\nu}_2}); \\
\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\mu}_i} &= \left.\frac{\partial^2 \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\mu}'_i}\right|_{\boldsymbol{\mu}_i=\mathbf{0}} = \Big[(\mathbf{I} - \mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}'_{p_i}\mathbf{I}^{(3)}_{p_i} + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i}\Big]\left(\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\right) \\
& \quad - \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i}\left(\mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\right); \\
\mathbf{D}^{(3)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\mu}} &= \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}}{\partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = \sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}} \otimes \mathbf{E}_{i,\mathbf{p}})\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\mu}_i}(\mathbf{I}_{\nu_3^2} \otimes \mathbf{E}'_{i,\boldsymbol{\nu}_2}); \\
\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\mu}_i} &= \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\mu}'_i}\right|_{\boldsymbol{\mu}_i=\mathbf{0}} \\
&= -\Big[(\mathbf{I} - \mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}'_{p_i}\mathbf{I}^{(3)}_{p_i} + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i}\Big]\Big[\left(\mathbf{I}_{(p_i,p_i)}\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi}} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\right) \\
& \quad + \left(\mathbf{I}_{(p_i,p_i)}\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\mu}_i} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}}\right)\mathbf{J}_{\nu_{2i},\nu_3}\Big] - \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i} \\
& \quad \times \Big[\left(\mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\right) + \left(\mathbf{I}_{(p_i,p_i)}\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\mu}_i} \otimes \mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\right)\mathbf{J}_{\nu_{2i},\nu_3}\Big] \\
& \quad + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(4)}_{p_i}\left(\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\varphi}} \otimes \mathbf{L}_{p_i}\mathbf{E}'_{i,\mathbf{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\right)\mathbf{J}^{*}_{\nu_{2i},\nu_3}; \\
\mathbf{D}^{(3)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\mu},\boldsymbol{\mu}} &= \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}}{\partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\mu}' \otimes \partial\,\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} \\
&= \sum_{i=1}^{k}(\mathbf{E}_{i,\mathbf{p}} \otimes \mathbf{E}_{i,\mathbf{p}})\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\mu}_i}(\mathbf{I}_{\nu_3} \otimes \mathbf{E}'_{i,\boldsymbol{\nu}_2} \otimes \mathbf{E}'_{i,\boldsymbol{\nu}_2}); \\
\mathbf{D}^{(3)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\mu}_i,\boldsymbol{\mu}_i} &= \left.\frac{\partial^3 \operatorname{vec}\mathbf{G}_i}{\partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\mu}'_i \otimes \partial\,\boldsymbol{\mu}'_i}\right|_{\boldsymbol{\mu}_i=\mathbf{0}} \\
&= -\Big[(\mathbf{I} - \mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}'_{p_i}\mathbf{I}^{(3)}_{p_i} + \mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}^{(2)}_{p_i,\boldsymbol{\lambda}_i}\mathbf{I}^{(3)}_{p_i,\boldsymbol{\lambda}_i}\Big]\left(\mathbf{I}_{(p_i,p_i)}\mathbf{D}^{(2)}_{\mathbf{g}_i:\boldsymbol{\varphi},\boldsymbol{\mu}_i} \otimes \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}\right)\mathbf{J}^{**}_{\nu_{2i},\nu_3}
\end{aligned}
$$

$$+\Big[(\mathbf{I}-\mathbf{H}_{\boldsymbol{\psi},i})\mathbf{A}_{1i}\mathbf{D}_{p_i}'\mathbf{I}_{p_i}^{(3)}+\mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}_{p_i,\boldsymbol{\lambda}_i}^{(2)}\mathbf{I}_{p_i,\boldsymbol{\lambda}_i}^{(3)}\Big]\left(\mathbf{D}_{\mathbf{g}_i:\boldsymbol{\varphi}}^{(1)}\otimes\mathbf{D}_{\mathbf{g}_i:\boldsymbol{\mu}_i,\boldsymbol{\mu}_i}^{(2)}\right)$$

$$+\mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}_{p_i,\boldsymbol{\lambda}_i}^{(2)}\mathbf{I}_{p_i}^{(4)}\left(\mathbf{D}_{\mathbf{g}_i:\boldsymbol{\mu}_i}^{(1)}\otimes\mathbf{L}_{p_i}\mathbf{E}_{i,\mathbf{p}}'\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)}\otimes\mathbf{D}_{\mathbf{g}_i:\boldsymbol{\mu}_i}^{(1)}\right)\left(\mathbf{I}_{(\nu_3,\nu_{2i})}\otimes\mathbf{I}_{\nu_{2i}}\right)$$

$$-\mathbf{H}_{\boldsymbol{\psi},i}\,\mathbf{I}_{p_i,\boldsymbol{\lambda}_i}^{(2)}\mathbf{I}_{p_i}^{(3)}\left(\mathbf{L}_{p_i}\mathbf{E}_{i,\mathbf{p}}'\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)}\otimes\mathbf{D}_{\mathbf{g}_i:\boldsymbol{\mu}_i,\boldsymbol{\mu}_i}^{(2)}\right);$$

where $\mathbf{D}_{p_i}$ *is the duplication matrix of order* $p_i^2\times p_i(p_i+1)/2$ *(Magnus and Neudecker, 1999, §3.8),* $\mathbf{N}_a^*$, $\mathbf{J}_a$, $\mathbf{J}_a^*$, $\mathbf{J}_{a,b}$, $\mathbf{J}_{a,b}^*$, $\mathbf{J}_{a,b}^{**}$, $\mathbf{I}_{p_i,\boldsymbol{\lambda}_i}^{(2)}$, $\mathbf{I}_{p_i}^{(3)}$, $\mathbf{I}_{p_i,\boldsymbol{\lambda}_i}^{(3)}$, *and* $\mathbf{I}_{p_i}^{(4)}$ *are defined in (34), and* $\mathbf{L}_{p_i}$ *is defined in* (37). □

6.4 Derivatives of Eigenvalues of $\boldsymbol{\Psi}$

The vector of eigenvalues of $\boldsymbol{\Psi}$ is parameterized as

$$\boldsymbol{\lambda}=w\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\},\ \text{where } w=\left(\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\}\right)^{-1}\tag{38}$$

and $\mathbf{T}_3$ and $\mathbf{T}_4$ are full column-rank design matrices with dimensions $\dot{p}\times q_3$ and $q_3\times q_4$ respectively. The design matrix $\mathbf{T}_4$ satisfies $\text{rank}\begin{pmatrix}\mathbf{T}_4 & \mathbf{1}_{q_3}\end{pmatrix}=q_4+1$. The q_4-vector $\boldsymbol{\xi}$ is parameterized as $\boldsymbol{\xi}=\boldsymbol{\xi}(\boldsymbol{\varphi})=\mathbf{V}_3\boldsymbol{\eta}+\mathbf{V}_4\boldsymbol{\varphi}$, where $\mathbf{V}_3$ and $\mathbf{V}_4$ are full column-rank matrices with dimensions $q_4\times c_2$ and $q_4\times(q_4-c_2)$, respectively; $\mathbf{V}_3'\mathbf{V}_4=\mathbf{0}$; $\boldsymbol{\eta}$ is an implicit function of $\boldsymbol{\varphi}$; $\boldsymbol{\xi}$ is subject to the constraints $\mathbf{C}_2'\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\}=\mathbf{0}$; and $c_2=\text{rank}(\mathbf{C}_2)$. If $\boldsymbol{\Psi}$ has a block diagonal structure with $k\geq 2$, then the eigenvalues within the i^{th} block must sum to p_i. That is, $\mathbf{1}_{p_i}'\boldsymbol{\lambda}_i=p_i$, where $\boldsymbol{\lambda}_i=\mathbf{E}_{i,\mathbf{p}}'\boldsymbol{\lambda}$ and $\mathbf{E}_{i,\mathbf{p}}$ is defined in (4). It is assumed that the matrix $\mathbf{C}_2$ incorporates these additional linear constraints on $\boldsymbol{\lambda}$.

Theorem 8. *Define* $\mathbf{W}_{\boldsymbol{\xi},1}$ *as* $\mathbf{W}_{\boldsymbol{\xi},1}=\text{Diag}\left(\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\}\right)\mathbf{T}_4$, *and express* $\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1}$ *in terms of its singular values and vectors;*

$$\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1}=\mathbf{U}_{\boldsymbol{\xi}}\begin{pmatrix}\mathbf{D}_{\boldsymbol{\xi}} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{A}_{\boldsymbol{\xi},1}'\\ \mathbf{A}_{\boldsymbol{\xi},2}'\end{pmatrix}=\mathbf{U}_{\boldsymbol{\xi}}\mathbf{D}_{\boldsymbol{\xi}}\mathbf{A}_{\boldsymbol{\xi},1}';\tag{39}$$

where $\mathbf{D}_{\boldsymbol{\xi}}$ *is an* $r_2\times r_2$ *diagonal matrix of nonzero singular values;* $r_2=\text{rank}(\mathbf{C}_2)$*;* $\mathbf{U}_{\boldsymbol{\xi}}\in\mathcal{O}_{r_2}$*;* $\mathbf{A}_{\boldsymbol{\xi},1}$ *has dimension* $q_4\times c_2$*;* $\mathbf{A}_{\boldsymbol{\xi},2}$ *has dimension* $q_4\times\nu_3$*;* $\nu_3=q_4-c_2$*; and* $\begin{pmatrix}\mathbf{A}_{\boldsymbol{\xi},1} & \mathbf{A}_{\boldsymbol{\xi},2}\end{pmatrix}\in\mathcal{O}_{q_4}$*. Let* $\mathbf{V}_3$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\boldsymbol{\xi},1})$*, where* $\mathcal{R}(\,\cdot\,)$ *denotes the vector space generated by the columns of a matrix. Similarly, let* $\mathbf{V}_4$ *be any matrix whose columns form a basis for* $\mathcal{R}(\mathbf{A}_{\boldsymbol{\xi},2})$*. This choice of* $\mathbf{V}^*=\begin{pmatrix}\mathbf{V}_3 & \mathbf{V}_4\end{pmatrix}$ *is delicate because* $\mathbf{V}^*$ *is defined as a function of* $\boldsymbol{\xi}$*, yet it will be treated as a constant when taking derivatives. For fixed* $\mathbf{V}^*$*, the first three derivatives of* $\boldsymbol{\lambda}$ *with respect to* $\boldsymbol{\varphi}'$ *are*

$$\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)}=\frac{\partial\,\boldsymbol{\lambda}}{\partial\,\boldsymbol{\varphi}'}=w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},1}\mathbf{V}_4,\quad w\text{ is defined in (38)};$$

$$\mathbf{H}_{\boldsymbol{\lambda}}=\mathbf{I}_{\dot{p}}-\frac{1}{\dot{p}}\boldsymbol{\lambda}\mathbf{1}_{\dot{p}}';$$

$$\begin{aligned}\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}}^{(2)}&=\frac{\partial^2\boldsymbol{\lambda}}{\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'}\\ &=-2w\left(\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)}\otimes\mathbf{1}_{\dot{p}}'\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},1}\mathbf{V}_3\right)\mathbf{N}_{\nu_3}+w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{H}_{\boldsymbol{\xi}}\mathbf{W}_{\boldsymbol{\xi},2}\left(\mathbf{V}_4\otimes\mathbf{V}_3\right);\end{aligned}$$

$$\begin{aligned}
\mathbf{H}_{\boldsymbol{\xi}} &= \mathbf{I}_{q_3} - \mathbf{W}_{\boldsymbol{\xi},1}(\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1})^{+}\mathbf{C}_2'; \\
(\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1})^{+} &= \mathbf{A}_{\boldsymbol{\xi},1}\mathbf{D}_{\boldsymbol{\xi}}^{-1}\mathbf{U}_{\boldsymbol{\xi}}' \textit{ is the Moore-Penrose inverse of } \mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1}; \\
\mathbf{W}_{\boldsymbol{\xi},2} &= \sum_{j=1}^{q_3} \mathbf{e}_j^{q_3} \exp\{\mathbf{e}_j^{q_3\prime}\mathbf{T}_4\boldsymbol{\xi}\}(\mathbf{e}_j^{q_3\prime}\mathbf{T}_4 \otimes \mathbf{e}_j^{q_3\prime}\mathbf{T}_4); \\
\mathbf{D}^{(3)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}} &= \frac{\partial^3\boldsymbol{\lambda}}{\partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\varphi}' \otimes \partial\,\boldsymbol{\varphi}'} \\
&= -w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{H}_{\boldsymbol{\xi}}\mathbf{W}_{\boldsymbol{\xi},2}\left[(\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1})^{+}\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},2}\,(\mathbf{V}_4 \otimes \mathbf{V}_4) \otimes \mathbf{V}_4\right]\mathbf{J}_{\nu_3} \\
&\quad -w\left(\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}} \otimes \mathbf{1}_{\dot{p}}'\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},1}\mathbf{V}_4\right)\mathbf{J}_{\nu_3} \\
&\quad -w\left[\mathbf{1}_{\dot{p}}'\mathbf{T}_3\mathbf{H}_{\boldsymbol{\xi}}\mathbf{W}_{\boldsymbol{\xi},2}\,(\mathbf{V}_4 \otimes \mathbf{V}_4) \otimes \mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\right]\mathbf{J}_{\nu_3} \\
&\quad +w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{H}_{\boldsymbol{\xi}}\mathbf{W}_{\boldsymbol{\xi},3}\,(\mathbf{V}_4 \otimes \mathbf{V}_4 \otimes \mathbf{V}_4)\,;\ \textit{and} \\
\mathbf{W}_{\boldsymbol{\xi},3} &= \sum_{j=1}^{q_3} \mathbf{e}_j^{q_3} \exp\{\mathbf{e}_j^{q_3\prime}\mathbf{T}_4\boldsymbol{\xi}\}(\mathbf{e}_j^{q_3\prime}\mathbf{T}_4 \otimes \mathbf{e}_j^{q_3\prime}\mathbf{T}_4 \otimes \mathbf{e}_j^{q_3\prime}\mathbf{T}_4).
\end{aligned}$$

Proof: For fixed $\mathbf{V}_3$ *and* $\mathbf{V}_4$*, the first derivative of* $\boldsymbol{\lambda}$ *with respect to* $\boldsymbol{\varphi}'$ *can be written as*

$$\begin{aligned}
\mathbf{D}_{\boldsymbol{\lambda},\boldsymbol{\varphi}} &= \frac{\partial}{\partial\boldsymbol{\epsilon}'}\left.\frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\left[\mathbf{V}_3\boldsymbol{\eta} + \mathbf{V}_4(\boldsymbol{\varphi}+\boldsymbol{\epsilon})\right]\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\left[\mathbf{V}_3\boldsymbol{\eta} + \mathbf{V}_4(\boldsymbol{\varphi}+\boldsymbol{\epsilon})\right]\}}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} \\
&= w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{W}_{\boldsymbol{\xi},1}\left(\mathbf{V}_3\left.\frac{\partial\,\boldsymbol{\eta}}{\partial\,\boldsymbol{\epsilon}'}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} + \mathbf{V}_4\right).
\end{aligned}$$

An expression for $\partial\,\boldsymbol{\eta}/\partial\,\boldsymbol{\epsilon}'$ *can be obtained by setting the derivative of the constraint in (5) with respect to* $\boldsymbol{\epsilon}$ *to zero and solving for* $\partial\,\boldsymbol{\eta}/\partial\,\boldsymbol{\epsilon}'$*. The result is*

$$\left.\frac{\partial}{\partial\,\boldsymbol{\epsilon}'}\,\mathbf{C}_2'\exp\{\odot\mathbf{T}_4\left[\mathbf{V}_3\boldsymbol{\eta} + \mathbf{V}_4(\boldsymbol{\varphi}+\boldsymbol{\epsilon})\right]\}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} = \mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1}\left(\mathbf{V}_3\left.\frac{\partial\,\boldsymbol{\eta}}{\partial\,\boldsymbol{\epsilon}'}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} + \mathbf{V}_4\right) = \mathbf{0}$$

$$\Longrightarrow \left.\frac{\partial\,\boldsymbol{\eta}}{\partial\,\boldsymbol{\epsilon}'}\right|_{\boldsymbol{\epsilon}=\mathbf{0}} = -\left(\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1}\mathbf{V}_3\right)^{-1}\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1}\mathbf{V}_4 = \mathbf{0}$$

because $\mathbf{C}_2'\mathbf{W}_{\boldsymbol{\xi},1} = \mathbf{U}_{\boldsymbol{\xi}}\mathbf{D}_{\boldsymbol{\xi}}\mathbf{A}_{\boldsymbol{\xi},1}'$ *and* $\mathbf{A}_{\boldsymbol{\xi},1}'\mathbf{V}_4 = \mathbf{0}$*. Therefore,* $\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}$ *simplifies to* $w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{W}_{\boldsymbol{\xi},1}\mathbf{V}_4$ *as claimed. The second and third derivatives are obtained in a parallel manner.* □

To reduce memory and computational requirements, matrix derivatives can be stored as sparse matrices. Also, to decrease the number of non-zero entries in the matrix derivatives, $\mathbf{V}_4$ can be equated to the transpose of the row reduced echelon form of $\mathbf{A}_{\boldsymbol{\xi},2}'$.

Corollary 2. If $\mathbf{C}_2$ in (5) has rank 0 (i.e., there are no additional constraints beyond $\mathbf{1}_{\dot{p}}'\boldsymbol{\lambda} = \dot{p}$), then $r_2 = 0$, $\mathbf{V}_4 = \mathbf{I}_{q_4}$, $\nu_3 = q_4$, $\boldsymbol{\varphi} = \boldsymbol{\xi}$ and the derivatives of $\boldsymbol{\lambda}$ simplify to

$$\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}} = w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},1},$$

$$
\begin{aligned}
\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}} &= -2w\left(\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}} \otimes \mathbf{1}'_{\dot{p}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},1}\right)\mathbf{N}_{\nu_3} + \dot{p}w\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},2};\\
\mathbf{D}^{(3)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}} &= -w\left(\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}} \otimes \mathbf{1}'_{\dot{p}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},1}\right)\mathbf{J}_{\nu_3} - w\left(\mathbf{1}'_{\dot{p}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},2} \otimes \mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\right)\mathbf{J}_{\nu_3}\\
&\quad + w\dot{p}\mathbf{H}_{\boldsymbol{\lambda}}\mathbf{T}_3\mathbf{W}_{\boldsymbol{\xi},3};
\end{aligned}
$$

where $\mathbf{N}_{\nu_3}$ and $\mathbf{J}_{\nu_3}$ are defined in (34). □

6.5 Derivatives of $\boldsymbol{\Sigma}$

6.5.1 *First-Order Derivatives of* $\operatorname{vec}\boldsymbol{\Sigma}$ *With Respect to* $\boldsymbol{\theta}$

Theorem 9. *Assume that $\boldsymbol{\theta}$ is partitioned as in* (3). *First-order derivatives of* $\operatorname{vec}\boldsymbol{\Sigma}$ *with respect to the components of $\boldsymbol{\theta}$, evaluated at $\boldsymbol{\mu} = \mathbf{0}$ are*

$$
\begin{aligned}
\left.\frac{\partial \operatorname{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= 2\mathbf{N}_{\dot{p}}(\boldsymbol{\sigma}_D\boldsymbol{\Psi} \otimes \mathbf{I}_{\dot{p}})\mathbf{L}_{\dot{p}}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}};\\
\left.\frac{\partial \operatorname{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= (\boldsymbol{\sigma}_D \otimes \boldsymbol{\sigma}_D)\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\mu}};\\
\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\mu}} &= \left.\frac{\partial \operatorname{vec}\boldsymbol{\Psi}}{\partial \boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot{p}}(\boldsymbol{\Gamma}\boldsymbol{\Lambda} \otimes \boldsymbol{\Gamma})\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}};\ \textit{and}\\
\left.\frac{\partial \operatorname{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= (\boldsymbol{\sigma}_D \otimes \boldsymbol{\sigma}_D)\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\varphi}};\\
\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\varphi}} &= \left.\frac{\partial \operatorname{vec}\boldsymbol{\Psi}}{\partial \boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot{p}}(\boldsymbol{\Gamma} \otimes \boldsymbol{\Gamma})\left[\mathbf{I}^{(2)}_{\dot{p},\boldsymbol{\lambda}}\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}} + \frac{1}{2}\mathbf{L}_{\dot{p}}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\right];
\end{aligned}
$$

where the derivatives of $\boldsymbol{\sigma}_d$, $\boldsymbol{\lambda}$, and $\mathbf{g}$ are given in Theorems 6, 8, and 7, respectively; $\mathbf{N}_{\dot{p}}$ *and* $\mathbf{I}^{(2)}_{\dot{p},\boldsymbol{\lambda}}$ *are defined in (34); and* $\mathbf{L}_{\dot{p}}$ *is defined in* (37).

6.5.2 *Second-Order Derivatives of* $\operatorname{vec}\boldsymbol{\Sigma}$ *With Respect to* $\boldsymbol{\theta}$

To obtain expressions for second- and third-order derivatives, the identities

$$
\begin{aligned}
(\mathbf{ABC} \otimes \mathbf{D})\mathbf{E} &= \left\{\mathbf{A} \otimes [\operatorname{vec}(\mathbf{C}')]' \otimes \mathbf{D}\right\}[\operatorname{vec}(\mathbf{B}') \otimes \mathbf{E}]\ \text{ and}\\
(\mathbf{A} \otimes \mathbf{BCD})\mathbf{E} &= [\mathbf{A} \otimes (\operatorname{vec}\mathbf{D})' \otimes \mathbf{B}]\,(\mathbf{E} \otimes \operatorname{vec}\mathbf{C}),
\end{aligned}
$$

are repeatedly used, where **A**, **B**, **C**, **D**, and **E** are any matrices that are conformable for multiplication.

Theorem 10. *Second-order derivatives of* $\operatorname{vec}\boldsymbol{\Sigma}$ *with respect to the components of $\boldsymbol{\theta}$, evaluated at $\boldsymbol{\mu} = \mathbf{0}$ are*

$$
\left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Sigma}}{\partial \boldsymbol{\tau}' \otimes \partial \boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot{p}}\Big[(\boldsymbol{\sigma}_D\boldsymbol{\Psi} \otimes \mathbf{I}_{\dot{p}})\mathbf{L}_{\dot{p}}\mathbf{D}^{(2)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau},\boldsymbol{\tau}}
$$

$$+(\mathbf{I}_{\dot p}\otimes \operatorname{vec}\boldsymbol{\Psi}\otimes \mathbf{I}_{\dot p})'(\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\otimes \mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}})\Big];$$

$$\left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Sigma}}{\partial\boldsymbol{\tau}'\otimes\partial\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\mathbf{I}_{\dot p}\otimes\boldsymbol{\sigma}_{D})\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\otimes\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\mu}}\right)$$

$$\left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Sigma}}{\partial\boldsymbol{\mu}'\otimes\partial\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = (\boldsymbol{\sigma}_{D}\otimes\boldsymbol{\sigma}_{D})\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\mu},\boldsymbol{\mu}};$$

$$\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\mu},\boldsymbol{\mu}} = \left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Psi}}{\partial\boldsymbol{\mu}'\otimes\partial\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\Gamma}\otimes\boldsymbol{\Gamma})\left[\mathbf{I}^{(2)}_{\dot p,\boldsymbol{\lambda}}\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\mu},\boldsymbol{\mu}} - \mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}(\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}})\right];$$

$$\left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = (\boldsymbol{\sigma}_{D}\otimes\boldsymbol{\sigma}_{D})\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\varphi},\boldsymbol{\varphi}};$$

$$\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\varphi},\boldsymbol{\varphi}} = \left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Psi}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\Gamma}\otimes\boldsymbol{\Gamma})\Big[2\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{L}_{pt}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\right)\mathbf{N}_{\nu_3}$$

$$-\mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}\left(\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\right)+\mathbf{I}^{(2)}_{\dot p,\boldsymbol{\lambda}}\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi}}+\frac{1}{2}\mathbf{L}_{\dot p}\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}}\Big];$$

$$\left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = (\boldsymbol{\sigma}_{D}\otimes\boldsymbol{\sigma}_{D})\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\varphi},\boldsymbol{\mu}};$$

$$\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\varphi},\boldsymbol{\mu}} = \left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Psi}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\Gamma}\otimes\boldsymbol{\Gamma})\Big[\mathbf{I}^{(2)}_{\dot p,\boldsymbol{\lambda}}\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\mu}} - \mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}\left(\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\right)$$

$$+\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\right)\Big];\ \textit{and}$$

$$\left.\frac{\partial^2 \operatorname{vec}\boldsymbol{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_{D}\otimes\mathbf{I}_{\dot p})\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\varphi}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\right);$$

where the derivatives of $\boldsymbol{\sigma}_d$, $\boldsymbol{\lambda}$, *and* $\mathbf{g}$ *are given in Theorems 6, 8, and 7, respectively;* $\mathbf{N}_{\dot p}$, $\mathbf{I}^{(2)}_{\dot p,\boldsymbol{\lambda}}$, $\mathbf{I}^{(3)}_{\dot p}$, *and* $\mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}$ *are defined in (34); and* $\mathbf{L}_{\dot p}$ *is defined in* (37).

6.5.3 *Third-Order Derivatives of* $\operatorname{vec}\boldsymbol{\Sigma}$ *With Respect to* $\boldsymbol{\theta}$

Theorem 11. *Third-order derivatives of* $\operatorname{vec}\boldsymbol{\Sigma}$ *with respect to the components of* $\boldsymbol{\theta}$, *evaluated at* $\boldsymbol{\mu}=\mathbf{0}$ *are*

$$\left.\frac{\partial^3 \operatorname{vec}\boldsymbol{\Sigma}}{\partial\boldsymbol{\tau}'\otimes\partial\boldsymbol{\tau}'\otimes\boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}\Big[(\mathbf{I}_{\dot p}\otimes\operatorname{vec}\boldsymbol{\Psi}\otimes\mathbf{I}_{\dot p})'\left(\mathbf{L}_{\dot p}\mathbf{D}^{(2)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau},\boldsymbol{\tau}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\right)\mathbf{J}_{\nu_1}$$

$$+(\boldsymbol{\sigma}_{D}\boldsymbol{\Psi}\otimes\mathbf{I}_{\dot p})\mathbf{L}_{\dot p}\mathbf{D}^{(3)}_{\boldsymbol{\sigma}:\boldsymbol{\tau},\boldsymbol{\tau},\boldsymbol{\tau}}\Big];$$

$$\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\mu}'\otimes\partial\boldsymbol{\mu}'\otimes\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_{D}\mathbf{\Gamma}\otimes\boldsymbol{\sigma}_{D}\mathbf{\Gamma})\Big\{\mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}\left(\mathbf{I}_{(\dot p,\dot p)}\,\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\mu},\boldsymbol{\mu}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\right)\mathbf{J}_{\nu_2}+\mathbf{I}^{(2)}_{\dot p,\boldsymbol{\lambda}}\mathbf{D}^{(3)}_{\mathbf{g}:\boldsymbol{\mu},\boldsymbol{\mu},\boldsymbol{\mu}}\Big\};$$

$$\begin{aligned}\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\varphi}'\otimes\boldsymbol{\varphi}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_{D}\mathbf{\Gamma}\otimes\boldsymbol{\sigma}_{D}\mathbf{\Gamma})\Big\{\Big[\mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}\left(\mathbf{I}_{(\dot p,\dot p)}\,\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\right)\\ &+\mathbf{I}^{(3)}_{\dot p}\Big[\left(\mathbf{L}_{\dot p}\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\right)+\left(\mathbf{I}_{(\dot p,\dot p)}\,\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\right)\Big]\Big]\mathbf{J}_{\nu_3}\\ &-\mathbf{I}^{(4)}_{\dot p}\left(\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\right)\mathbf{J}^{*}_{\nu_3}+\mathbf{I}^{(2)}_{\dot p,\boldsymbol{\lambda}}\mathbf{D}^{(3)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}}+\frac{1}{2}\mathbf{L}_{\dot p}\mathbf{D}^{(3)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}}\Big\}\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\mu}'\otimes\partial\boldsymbol{\tau}'\otimes\boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= 2\mathbf{N}_{\dot p}\Big\{\mathbf{I}^{(4)}_{\dot p}\left(\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\otimes\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\mu}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\right)\left(\mathbf{I}_{(\nu_2,\nu_1)}\otimes\mathbf{I}_{\nu_1}\right)\\ &+(\boldsymbol{\sigma}_{D}\otimes\mathbf{I}_{\dot p})\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\mu}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(2)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau},\boldsymbol{\tau}}\right)\Big\};\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\tau}'\otimes\boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= 2\mathbf{N}_{\dot p}\Big\{\mathbf{I}^{(4)}_{\dot p}\left(\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\otimes\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\varphi}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\right)\left(\mathbf{I}_{(\nu_3,\nu_1)}\otimes\mathbf{I}_{\nu_1}\right)\\ &+(\boldsymbol{\sigma}_{D}\otimes\mathbf{I}_{\dot p})\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{D}^{(1)}_{\boldsymbol{\psi}:\boldsymbol{\varphi}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(2)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau},\boldsymbol{\tau}}\right)\Big\};\end{aligned}$$

$$\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\mu}'\otimes\partial\boldsymbol{\mu}'\otimes\boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_{D}\otimes\mathbf{I}_{\dot p})\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\mu},\boldsymbol{\mu}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\right);$$

$$\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\varphi}'\otimes\boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_{D}\otimes\mathbf{I}_{\dot p})\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{D}^{(2)}_{\boldsymbol{\psi}:\boldsymbol{\varphi},\boldsymbol{\varphi}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}\right);$$

$$\begin{aligned}\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\varphi}'\otimes\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_{D}\mathbf{\Gamma}\otimes\boldsymbol{\sigma}_{D}\mathbf{\Gamma})\Big\{\mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}\Big[\left(\mathbf{I}_{(\dot p,\dot p)}\,\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\right)\\ &+\left(\mathbf{I}_{(\dot p,\dot p)}\,\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\mu}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\right)\mathbf{J}_{\nu_2,\nu_3}\Big]\\ &+\mathbf{I}^{(3)}_{\dot p}\Big[\left(\mathbf{L}_{\dot p}\mathbf{D}^{(2)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi},\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\right)+\left(\mathbf{I}_{(\dot p,\dot p)}\,\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\mu}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\right)\mathbf{J}_{\nu_2,\nu_3}\Big]\\ &-\mathbf{I}^{(4)}_{\dot p}\left(\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\otimes\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\right)\mathbf{J}^{*}_{\nu_2,\nu_3}+\mathbf{I}^{(2)}_{\dot p,\boldsymbol{\lambda}}\mathbf{D}^{(3)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\mu}}\Big\};\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial^3 \operatorname{vec}\mathbf{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\mu}'\otimes\boldsymbol{\mu}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} &= 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_{D}\mathbf{\Gamma}\otimes\boldsymbol{\sigma}_{D}\mathbf{\Gamma})\Big\{\mathbf{I}^{(3)}_{\dot p}\left(\mathbf{L}_{\dot p}\mathbf{D}^{(1)}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\mu},\boldsymbol{\mu}}\right)\\ &-\mathbf{I}^{(3)}_{\dot p,\boldsymbol{\lambda}}\Big[\left(\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\varphi}}\otimes\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\mu},\boldsymbol{\mu}}\right)-\left(\mathbf{I}_{(\dot p,\dot p)}\,\mathbf{D}^{(2)}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\mu}}\otimes\mathbf{D}^{(1)}_{\mathbf{g}:\boldsymbol{\mu}}\right)\mathbf{J}^{**}_{\nu_2,\nu_3}\Big]\end{aligned}$$

$$-\mathbf{I}_{\dot p}^{(4)}\left(\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)}\otimes\mathbf{L}_{\dot p}\mathbf{D}_{\boldsymbol{\lambda}:\boldsymbol{\varphi}}^{(1)}\otimes\mathbf{D}_{\mathbf{g}:\boldsymbol{\mu}}^{(1)}\right)\left(\mathbf{I}_{(\nu_3,\nu_2)}\otimes\mathbf{I}_{\nu_2}\right)+\mathbf{I}_{\dot p,\boldsymbol{\lambda}}^{(2)}\mathbf{D}_{\mathbf{g}:\boldsymbol{\varphi},\boldsymbol{\mu},\boldsymbol{\mu}}^{(3)}\Bigg\};$$

$$\left.\frac{\partial^3\operatorname{vec}\boldsymbol{\Sigma}}{\partial\boldsymbol{\varphi}'\otimes\partial\boldsymbol{\mu}'\otimes\boldsymbol{\tau}'}\right|_{\boldsymbol{\mu}=\mathbf{0}} = 2\mathbf{N}_{\dot p}(\boldsymbol{\sigma}_D\otimes\mathbf{I}_{\dot p})\mathbf{I}_{\dot p}^{(3)}\left(\mathbf{D}_{\boldsymbol{\psi}:\boldsymbol{\varphi},\boldsymbol{\mu}}^{(2)}\otimes\mathbf{L}_{\dot p}\mathbf{D}_{\boldsymbol{\sigma}_d:\boldsymbol{\tau}}^{(1)}\right);$$

where the derivatives of $\boldsymbol{\sigma}_d$*,* $\boldsymbol{\lambda}$*, and* $\mathbf{g}$ *are given in Theorems 6, 8, and 7, respectively;* $\mathbf{N}_{\dot p}$*,* $\mathbf{J}_a$*,* $\mathbf{J}_{a,b}$*,* $\mathbf{J}_{a,b}^{*}$*,* $\mathbf{J}_{a,b}^{**}$*,* $\mathbf{I}_{\dot p,\boldsymbol{\lambda}}^{(2)}$*,* $\mathbf{I}_{\dot p}^{(3)}$*,* $\mathbf{I}_{\dot p,\boldsymbol{\lambda}}^{(3)}$*, and* $\mathbf{I}_{\dot p}^{(4)}$ *are defined in (34); and* $\mathbf{L}_{\dot p}$ *is defined in* (37).

7 Initial Guesses

7.1 Guess for $\boldsymbol{\tau}$

Denote the p-vector of sample standard deviations by $\mathbf{s}_d$ and denote the diagonal matrix of standard deviations by $\mathbf{s}_D$. Using the delta method, it is readily shown that if $\mathbf{S}$ has a Wishart distribution, then

$$\sqrt{n}(\mathbf{s}_d-\boldsymbol{\sigma}_d)\xrightarrow{\text{dist}}\mathrm{N}\left(\mathbf{0},2^{-1}{\boldsymbol{\sigma}_D}^{-1}\boldsymbol{\Sigma}^{\odot 2}{\boldsymbol{\sigma}_D}^{-1}\right). \tag{40}$$

If $\mathbf{S}$ is not Wishart, then the asymptotic variance of $\sqrt{n}(\mathbf{s}_d-\boldsymbol{\sigma}_d)$ depends on the kurtosis of the parent distribution that generated $\mathbf{Y}$. Nonetheless, for purposes of obtaining initial guesses, there is little harm in using the asymptotic distribution in (40).

If possible, a two stage generalized least squares procedure is used to compute an initial guess for $\boldsymbol{\tau}$. In the first stage, an estimate of $\exp\{\odot\mathbf{T}_2\boldsymbol{\tau}\}$ is computed. If all elements of the estimate are positive, then an estimate for $\boldsymbol{\tau}$ is computed in the second stage. The first stage estimate is

$$\exp\{\odot\mathbf{T}_2\widetilde{\boldsymbol{\tau}}\}=\left[\mathbf{T}_1'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D\mathbf{T}_1\right]^{-1}\mathbf{T}_1'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D\mathbf{s}_d.$$

Using the delta method, it is readily shown that

$$\operatorname{Var}\ln\left\{\odot\left[\mathbf{T}_1'\boldsymbol{\sigma}_D\left(\boldsymbol{\Sigma}^{\odot 2}\right)^{-1}\boldsymbol{\sigma}_D\mathbf{T}_1\right]^{-1}\mathbf{T}_1'\boldsymbol{\sigma}_D\left(\boldsymbol{\Sigma}^{\odot 2}\right)^{-1}\boldsymbol{\sigma}_D\mathbf{s}_d\right\}\approx$$

$$\operatorname{Diag}\left(\exp\{\odot(-\mathbf{T}_2\boldsymbol{\tau})\}\right)\left[\mathbf{T}_1'\boldsymbol{\sigma}_D\left(\boldsymbol{\Sigma}^{\odot 2}\right)^{-1}\boldsymbol{\sigma}_D\mathbf{T}_1\right]^{-1}\operatorname{Diag}\left(\exp\{\odot(-\mathbf{T}_2\boldsymbol{\tau})\}\right).$$

The second stage estimate is

$$\widehat{\boldsymbol{\tau}}=\left[\mathbf{T}_2'\operatorname{Diag}\left(\exp\{\odot\mathbf{T}_2\widetilde{\boldsymbol{\tau}}\}\right)\mathbf{T}_1'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D\mathbf{T}_1\operatorname{Diag}(\exp\{\odot\mathbf{T}_2\widetilde{\boldsymbol{\tau}}\})\mathbf{T}_2\right]^{-1}$$

$$\times\,\mathbf{T}_2'\operatorname{Diag}\left(\exp\{\odot\mathbf{T}_2\widetilde{\boldsymbol{\tau}}\}\right)\mathbf{T}_1'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D\mathbf{T}_1\operatorname{Diag}(\exp\{\odot\mathbf{T}_2\widetilde{\boldsymbol{\tau}}\})\ln\left[\odot\exp\{\odot\mathbf{T}_2\widetilde{\boldsymbol{\tau}}\}\right].$$

If any component of the first stage estimate is negative, then an alternative two stage method is used. The first stage of the alternative method employs the Gauss-Newton algorithm to minimize

$$\mathit{SSE}(\omega)=(\mathbf{s}_d-\mathbf{T}_1\exp\{\odot\mathbf{T}_2\mathbf{1}_{\nu_1}\omega\})'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D(\mathbf{s}_d-\mathbf{T}_1\exp\{\odot\mathbf{T}_2\mathbf{1}_{\nu_1}\omega\})$$

with respect to the scalar ω, where the initial guess for ω is zero. At iteration $h+1$, the estimate of ω is updated by

$$\widehat{\omega}_{h+1} = \widehat{\omega}_h+\alpha_h\left[\mathbf{x}'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D\mathbf{x}\right]^{-1}\mathbf{x}'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D(\mathbf{s}_d-\widehat{\boldsymbol{\sigma}}_d),\ \text{where}$$

$$\begin{aligned}\mathbf{x} &= \mathbf{T}_1 \operatorname{Diag}\left(\exp\{\odot\mathbf{T}_1\mathbf{1}_{\nu_1}\widehat{\omega}_h\}\right)\mathbf{T}_2\mathbf{1}_{\nu_1},\\ \widehat{\boldsymbol{\sigma}}_d &= \mathbf{T}_2\exp\{\odot\mathbf{T}_2\mathbf{1}_{\nu_1}\widehat{\omega}_h\},\end{aligned}$$

and $\alpha_h \in (0, 1]$.

The second stage of the alternative method employs the Gauss-Newton algorithm to minimize

$$SSE(\boldsymbol{\tau}) = (\mathbf{s}_d - \mathbf{T}_1\exp\{\odot\mathbf{T}_2\boldsymbol{\tau}\})'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D(\mathbf{s}_d - \mathbf{T}_1\exp\{\odot\mathbf{T}_2\boldsymbol{\tau}\})$$

with respect to $\boldsymbol{\tau}$, where the initial guess for $\boldsymbol{\tau}$ is $\mathbf{1}_{\nu_1}\widehat{\omega}$ and $\widehat{\omega}$ is the minimizer computed in stage one. At iteration $h+1$, the estimate of $\boldsymbol{\tau}$ is updated by

$$\begin{aligned}\widehat{\boldsymbol{\tau}}_{h+1} &= \widehat{\boldsymbol{\tau}}_h + \alpha_h\left[\mathbf{X}'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D\mathbf{X}\right]^{-1}\mathbf{X}'\mathbf{s}_D\left(\mathbf{S}^{\odot 2}\right)^{-1}\mathbf{s}_D(\mathbf{s}_d - \widehat{\boldsymbol{\sigma}}_d), \text{ where}\\ \mathbf{X} &= \mathbf{T}_1\operatorname{Diag}\left(\exp\{\odot\mathbf{T}_1\widehat{\boldsymbol{\tau}}_h\}\right)\mathbf{T}_2,\\ \widehat{\boldsymbol{\sigma}}_d &= \mathbf{T}_2\exp\{\odot\mathbf{T}_2\widehat{\boldsymbol{\tau}}_h\},\end{aligned}$$

and $\alpha_h \in (0, 1]$.

7.2 Guesses for $\boldsymbol{\xi}$ and $\boldsymbol{\varphi}$

The strategy for obtaining initial guesses for the eigenvalue parameters is first to obtain an initial guess for $\boldsymbol{\xi}$. The initial guess for $\boldsymbol{\xi}$ is then used as the starting point for obtaining an initial guess for $\boldsymbol{\varphi}$.

7.2.1 *Initial Guess for* $\boldsymbol{\xi}$

Denote the $\dot{p}$-vector of eigenvalues of $\mathbf{R}$ as $\boldsymbol{\ell}$ and denote the diagonal matrix of sample eigenvalues by $\boldsymbol{\ell}_D$. It can be shown (Kollo and Neudecker, 1993; Boik, 1998) that if $\mathbf{S}$ has a Wishart distribution, then

$$\sqrt{n}(\boldsymbol{\ell} - \boldsymbol{\lambda}) \xrightarrow{\text{dist}} \mathrm{N}\left(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\lambda}}\right) \text{ where} \tag{41}$$

$$\boldsymbol{\Omega}_{\boldsymbol{\lambda}} = 2\mathbf{L}_{\dot{p}}'(\boldsymbol{\Gamma}\otimes\boldsymbol{\Gamma})'\left[\mathbf{I}_{\dot{p}^2} - (\boldsymbol{\Psi}\otimes\mathbf{I}_{\dot{p}})\mathbf{W}_{\dot{p}}\right](\boldsymbol{\Psi}\otimes\boldsymbol{\Psi})\left[\mathbf{I}_{\dot{p}^2} - \mathbf{W}_{\dot{p}}(\boldsymbol{\Psi}\otimes\mathbf{I}_{\dot{p}})\right](\boldsymbol{\Gamma}\otimes\boldsymbol{\Gamma})\mathbf{L}_{\dot{p}},$$

$$\mathbf{W}_{\dot{p}} = \sum_{i=1}^{\dot{p}}(\mathbf{e}_i^{\dot{p}}\mathbf{e}_i^{\dot{p}\prime}\otimes\mathbf{e}_i^{\dot{p}}\mathbf{e}_i^{\dot{p}\prime}),$$

and $\mathbf{L}_{\dot{p}}$ is defined in (37). Note that $\Pr(\mathbf{1}_{\dot{p}}'\boldsymbol{\ell} = \dot{p}) = 1$ and, therefore, $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}\mathbf{1}_{\dot{p}} = \mathbf{0}$.

A two-stage procedure is used to obtain an initial guess for ξ. The first stage employs the Gauss-Newton algorithm to minimize

$$SSE(\omega) = \left(\boldsymbol{\ell} - \frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\omega\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\omega\}}\right)'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'\left(\boldsymbol{\ell} - \frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\omega\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\omega\}}\right)$$

with respect to the scalar ω, where the initial guess for ω is zero, $\mathbf{K}$ is a $\dot{p}\times(\dot{p}-1)$ matrix whose columns form a basis for the null space of $\mathbf{1}_{\dot{p}}'$, and $\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}$ is a consistent estimator of $\boldsymbol{\Omega}_{\boldsymbol{\lambda}}$. At iteration $h+1$, the estimate is updated by

$$\widehat{\omega}_{h+1} = \widehat{\omega}_h + \alpha_h\left\{\mathbf{x}'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'\mathbf{x}\right\}^{-1}\mathbf{x}'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'(\boldsymbol{\ell} - \widehat{\boldsymbol{\lambda}}_h), \text{ where}$$

$$\begin{aligned}\widehat{\boldsymbol{\lambda}}_h &= \frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\widehat{\omega}_h\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\widehat{\omega}_h\}},\\ \mathbf{x} &= \frac{\dot{p}(\mathbf{I}_{\dot{p}}-\widehat{\boldsymbol{\lambda}}_h\dot{p}^{-1}\mathbf{1}_{\dot{p}}')\mathbf{T}_3\,\mathrm{Diag}\left(\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\widehat{\omega}_h\}\right)\mathbf{T}_4\mathbf{1}_{q_4}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\mathbf{1}_{q_4}\widehat{\omega}_h\}},\end{aligned}$$

and $\alpha_h \in (0,1]$.

The second stage employs the Gauss-Newton algorithm to minimize

$$\mathit{SSE}(\boldsymbol{\xi}) = \left(\boldsymbol{\ell}-\frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\}}\right)'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'\left(\boldsymbol{\ell}-\frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}\}}\right)$$

with respect to $\boldsymbol{\xi}$, where the initial guess for $\boldsymbol{\xi}$ is $\mathbf{1}_{q_4}\widehat{\omega}$ and $\widehat{\omega}$ is the minimizer from stage one. At iteration $h+1$, the estimate is updated by

$$\begin{aligned}\widehat{\boldsymbol{\xi}}_{h+1} &= \widehat{\boldsymbol{\xi}}_h+\alpha_h\left\{\mathbf{X}'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'\mathbf{X}\right\}^{-1}\mathbf{X}'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'(\boldsymbol{\ell}-\widehat{\boldsymbol{\lambda}}_h),\ \text{where}\\ \widehat{\boldsymbol{\lambda}}_h &= \frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\widehat{\boldsymbol{\xi}}_h\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\widehat{\boldsymbol{\xi}}_h\}},\\ \mathbf{X} &= \frac{\dot{p}(\mathbf{I}_{\dot{p}}-\widehat{\boldsymbol{\lambda}}_h\dot{p}^{-1}\mathbf{1}_{\dot{p}}')\mathbf{T}_3\,\mathrm{Diag}\left(\exp\{\odot\mathbf{T}_4\widehat{\boldsymbol{\xi}}_h\}\right)\mathbf{T}_4}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\widehat{\boldsymbol{\xi}}_h\}},\end{aligned}$$

and $\alpha_h \in (0,1]$.

7.2.2 *Initial Guess for* $\boldsymbol{\varphi}$

If $\boldsymbol{\lambda}$ is not subject to linear constraints other than $\mathbf{1}_{\dot{p}}'\boldsymbol{\lambda}=\dot{p}$, then the initial guess for $\boldsymbol{\varphi}$ is $\widehat{\boldsymbol{\xi}}$, where $\widehat{\boldsymbol{\xi}}$ is the initial guess for $\boldsymbol{\xi}$.

Suppose that $\boldsymbol{\lambda}$ is subject to the constraints $\mathbf{C}_1'\boldsymbol{\lambda}=\mathbf{c}_0$ in addition to $\mathbf{1}_{\dot{p}}'\boldsymbol{\lambda}=\dot{p}$. In this case, the guess for $\boldsymbol{\varphi}$ is the minimizer of

$$\mathit{SSE}(\boldsymbol{\varphi}) = \left(\boldsymbol{\ell}-\frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}(\boldsymbol{\varphi})\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}(\boldsymbol{\varphi})\}}\right)'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'\left(\boldsymbol{\ell}-\frac{\dot{p}\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}(\boldsymbol{\varphi})\}}{\mathbf{1}_{\dot{p}}'\mathbf{T}_3\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}(\boldsymbol{\varphi})\}}\right)$$

with respect to $\boldsymbol{\varphi}$, where $\xi(\boldsymbol{\varphi})$ satisfies $\mathbf{C}_2'\exp\{\odot\mathbf{T}_4\boldsymbol{\xi}(\boldsymbol{\varphi})\}=\mathbf{0}$ and $\mathbf{C}_2$ is a $q_3\times r_2$ matrix whose columns form a basis for $\mathcal{R}\left[\mathbf{T}_3'(\mathbf{C}_1-\mathbf{1}_{\dot{p}}\dot{p}^{-1}\mathbf{c}_0')\right]$. See equation (16). A Gauss-Newton algorithm is used to minimize $\mathit{SSE}(\boldsymbol{\varphi})$. Set $\widehat{\boldsymbol{\xi}}_0=\widehat{\boldsymbol{\xi}}$, where $\widehat{\boldsymbol{\xi}}$ is the initial guess for $\boldsymbol{\xi}$. Define $\mathbf{W}_{\hat{\boldsymbol{\xi}}_h,1}$ as

$$\mathbf{W}_{\hat{\boldsymbol{\xi}}_h,1}\overset{\text{def}}{=}\mathrm{diag}\left(\exp\{\odot\mathbf{T}_4\widehat{\boldsymbol{\xi}}_h\}\right)\mathbf{T}_4.$$

and denote the singular value decomposition of $\mathbf{C}_2'\mathbf{W}_{\hat{\boldsymbol{\xi}}_h,1}$

$$\mathbf{C}_2'\mathbf{W}_{\hat{\boldsymbol{\xi}}_h,1}=\mathbf{U}_{\hat{\boldsymbol{\xi}}_h}\begin{pmatrix}\mathbf{D}_{\hat{\boldsymbol{\xi}}_h} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{A}_{\hat{\boldsymbol{\xi}},1}'\\ \mathbf{A}_{\hat{\boldsymbol{\xi}},2}'\end{pmatrix},$$

where $\mathbf{D}_{\hat{\boldsymbol{\xi}}_h}$ is an $r_2\times r_2$ diagonal matrix of non-zero singular values, $\mathbf{U}_{\hat{\boldsymbol{\xi}}_h}\in\mathcal{O}_{r_2}$, and $\mathbf{A}_{\hat{\boldsymbol{\xi}}}=\begin{pmatrix}\mathbf{A}_{\hat{\boldsymbol{\xi}},1} & \mathbf{A}_{\boldsymbol{\xi},2}\end{pmatrix}\in\mathcal{O}_{q_4}$. Write $\widehat{\boldsymbol{\xi}}_h$ as

$$\widehat{\boldsymbol{\xi}}_h=\mathbf{V}_3\widehat{\boldsymbol{\eta}}_h+\mathbf{V}_4\widehat{\boldsymbol{\varphi}}_h,\ \text{where}\ \widehat{\boldsymbol{\eta}}_h=(\mathbf{V}_3'\mathbf{V}_3)^{-1}\mathbf{V}_3'\widehat{\boldsymbol{\xi}}_h,\quad \widehat{\boldsymbol{\varphi}}_h=(\mathbf{V}_4'\mathbf{V}_4)^{-1}\mathbf{V}_4'\widehat{\boldsymbol{\xi}}_h,$$

$\mathbf{V}_3$ is a $q_4 \times r_2$ matrix whose columns form a basis for $\mathcal{R}(\mathbf{A}_{\hat{\boldsymbol{\xi}},1})$, and $\mathbf{V}_4$ is a $q_4 \times \nu_3$ matrix whose columns form a basis for $\mathcal{R}(\mathbf{A}_{\hat{\boldsymbol{\xi}},2})$. At iteration $h+1$, the estimate of $\boldsymbol{\varphi}$ is updated as

$$\begin{aligned}
\widehat{\boldsymbol{\varphi}}_{h+1} &= \widehat{\boldsymbol{\varphi}}_h + \alpha_h \left\{ \mathbf{X}'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'\mathbf{X} \right\}^{-1} \mathbf{X}'\mathbf{K}\left(\mathbf{K}'\boldsymbol{\Omega}_{\hat{\boldsymbol{\lambda}}}\mathbf{K}\right)^{-1}\mathbf{K}'(\boldsymbol{\ell} - \widehat{\boldsymbol{\lambda}}_h), \text{ where} \\
\widehat{\boldsymbol{\lambda}}_h &= \frac{\dot{p}\mathbf{T}_3 \exp\{\odot \mathbf{T}_4 \widehat{\boldsymbol{\xi}}_h\}}{\mathbf{1}'_{\dot{p}}\mathbf{T}_3 \exp\{\odot \mathbf{T}_4 \widehat{\boldsymbol{\xi}}_h\}}, \\
\mathbf{X} &= \frac{\dot{p}(\mathbf{I}_{\dot{p}} - \widehat{\boldsymbol{\lambda}}_h \dot{p}^{-1}\mathbf{1}'_{\dot{p}})\mathbf{T}_3 \mathbf{W}_{\hat{\boldsymbol{\xi}}_h}\mathbf{V}_4}{\mathbf{1}'_{\dot{p}}\mathbf{T}_3 \exp\{\odot \mathbf{T}_4 \widehat{\boldsymbol{\xi}}_h\}},
\end{aligned}$$

and $\alpha_h \in (0,1]$. Given $\widehat{\boldsymbol{\varphi}}_{h+1}$, the estimate of the corresponding implicit parameter, $\widehat{\boldsymbol{\eta}}_{h+1}$, is obtained by solving

$$\mathbf{C}'_2 \exp\{\odot \mathbf{T}_4 \left(\mathbf{V}_3 \widehat{\boldsymbol{\eta}}_{h+1} + \mathbf{V}_4 \widehat{\boldsymbol{\varphi}}_{h+1}\right)\} = \mathbf{0}.$$

Set $\widehat{\boldsymbol{\eta}}_{h+1,0} = \widehat{\boldsymbol{\eta}}_h$. The modified Newton update for $\widehat{\boldsymbol{\eta}}_{h+1}$ is

$$\begin{aligned}
\widehat{\boldsymbol{\eta}}_{h+1,j+1} &= \widehat{\boldsymbol{\eta}}_{h+1,j} - \alpha_{h+1,j} \left\{ \mathbf{C}'_2 \operatorname{Diag}\left[\exp\left\{\odot \mathbf{T}_4 \left(\mathbf{V}_3 \widehat{\boldsymbol{\eta}}_{h+1,j} + \mathbf{V}_4 \widehat{\boldsymbol{\varphi}}_{h+1}\right)\right\}\right] \mathbf{T}_4 \mathbf{V}_3 \right\}^{-1} \\
&\quad \times \mathbf{C}'_2 \exp\left\{\odot \mathbf{T}_4 \left(\mathbf{V}_3 \widehat{\boldsymbol{\eta}}_{h+1,j} + \mathbf{V}_4 \widehat{\boldsymbol{\varphi}}_{h+1}\right)\right\},
\end{aligned}$$

where $\alpha_{h+1,j} \in (0,1]$.

7.3 Guess for $\boldsymbol{\Gamma}$

A two step procedure is used to obtain an initial guess for $\boldsymbol{\Gamma}$. In step 1, an estimate that satisfies $\mathbf{e}_i^{\dot{p}\prime}\widehat{\boldsymbol{\Gamma}}\widehat{\boldsymbol{\Lambda}}\widehat{\boldsymbol{\Gamma}}\mathbf{e}_i^{\dot{p}} = 1$ for $i = 1, \ldots, \dot{p}$ is obtained, where $\widehat{\boldsymbol{\Lambda}}$ is a diagonal matrix with diagonal entries equal to the estimates computed in §7.2.2. In step 2, the initial estimate is improved.

7.3.1 *Computing an Initial Estimate*

Denote the i^{th} diagonal block of the sample correlation matrix by $\mathbf{R}_i$; i.e., $\mathbf{R}_i = \mathbf{E}'_{i,\mathbf{p}}\mathbf{R}\mathbf{E}_{i,\mathbf{p}}$, for $i = 1, \ldots, k$. The $p_i(p_i-1)/2 \times 1$ vector of distinct entries of $\mathbf{R}_i$ below the diagonal can be obtained as $\mathbf{H}'_i \operatorname{vec}(\mathbf{R}_i)$, where

$$\mathbf{H}_i = \mathbf{N}_{p_i} \sum_{s=1}^{p_i-1} \sum_{t=s+1}^{p_i} (\mathbf{e}_s^{p_i} \otimes \mathbf{e}_t^{p_i})\mathbf{e}_u^{p_i(p_i-1)/2\prime},$$

$u = (2p_i - 1 - s)s/2 + t - p_i$, and $\mathbf{N}_{p_i}$ is defined in (34). Using the delta method, it is readily shown that if $\mathbf{S}$ has a Wishart distribution, then

$$\begin{aligned}
&\sqrt{n}\mathbf{H}'_i \operatorname{vec}(\mathbf{R}_i - \boldsymbol{\Psi}_i) \xrightarrow{\text{dist}} \mathrm{N}(\mathbf{0}, \mathbf{M}_i), \text{ where} \\
&\mathbf{M}_i = 2\mathbf{H}'_i \left[\mathbf{I}_{p_i^2} - (\boldsymbol{\Psi}_i \otimes \mathbf{I}_{p_i})\mathbf{W}_{p_i}\right] (\boldsymbol{\Psi}_i \otimes \boldsymbol{\Psi}_i) \left[\mathbf{I}_{p_i^2} - \mathbf{W}_{p_i}(\boldsymbol{\Psi}_i \otimes \mathbf{I}_{p_i})\right] \mathbf{H}_i,
\end{aligned}$$

and $\mathbf{W}_{p_i}$ is defined in (41).

Denote the $p_i \times p_i$ matrix containing the eigenvectors of $\mathbf{R}_i$ by $\widetilde{\boldsymbol{\Gamma}}_i$. An initial guess for $\boldsymbol{\Gamma}_i$ can be obtained by employing a modification of algorithm AS 213 (Lin and Bendel, 1985). The modifications are as follows.

1. Set the eigenvalues of $\widehat{\boldsymbol{\Psi}}_i$ to
$$\widehat{\boldsymbol{\lambda}}_i = \mathbf{E}'_{i,\mathbf{p}} \frac{\dot{p}\mathbf{T}_3 \exp\{\odot \mathbf{T}_4 \widehat{\boldsymbol{\xi}}\}}{\mathbf{1}'_{\dot{p}} \mathbf{T}_3 \exp\{\odot \mathbf{T}_4 \widehat{\boldsymbol{\xi}}\}},$$
and set $\widehat{\boldsymbol{\Lambda}}_i = \operatorname{Diag}(\widehat{\boldsymbol{\lambda}}_i)$, where $\widehat{\boldsymbol{\xi}}$ is the initial guess for $\boldsymbol{\xi}$ computed in §7.2.2.

2. Denote the estimate of $\boldsymbol{\Gamma}_i$ at iteration h by $\widehat{\boldsymbol{\Gamma}}_{i,h}$ and set the initial estimate to $\widehat{\boldsymbol{\Gamma}}_{i,0} = \widetilde{\boldsymbol{\Gamma}}_i$. It is assumed that the columns of $\widetilde{\boldsymbol{\Gamma}}_i$ have been ordered to correspond to the sorted eigenvalues of $\mathbf{R}_i$ and that the components of $\widehat{\boldsymbol{\lambda}}_i$ are ordered from largest to smallest. The latter ordering is accomplished by means of the $\mathbf{T}_3$ design matrix.

3. At iteration h of algorithm AS 213 choose the largest and smallest diagonal elements of $\widehat{\boldsymbol{\Psi}}_i = \widehat{\boldsymbol{\Gamma}}_{i,h} \widehat{\boldsymbol{\Lambda}}_i \widehat{\boldsymbol{\Gamma}}'_{i,h}$.

4. The update from $\widehat{\boldsymbol{\Gamma}}_{i,h}$ to $\widehat{\boldsymbol{\Gamma}}_{i,h+1}$ is obtained by multiplication by an elementary orthogonal matrix. This multiplication sets one of the two selected diagonal entries of the correlation matrix to one. For each of the two diagonal elements, in turn, update the estimate to $\boldsymbol{\Gamma}_{i,h+1}$ and choose the update that has the smaller
$$\mathit{SSE} = \left[\mathbf{H}'_i \operatorname{vec}(\mathbf{R}_i - \widehat{\boldsymbol{\Psi}}_{i,h+1})\right]' \widehat{\mathbf{M}}^{-1} \mathbf{H}'_i \operatorname{vec}(\mathbf{R}_i - \widehat{\boldsymbol{\Psi}}_{i,h+1}),$$
where $\widehat{\boldsymbol{\Psi}}_{i,h+1} = \widehat{\boldsymbol{\Gamma}}_{i,h+1} \widehat{\boldsymbol{\Lambda}}_i \widehat{\boldsymbol{\Gamma}}'_{i,h+1}$ and $\widehat{\mathbf{M}}_i$ is a consistent estimator of $\mathbf{M}_i$.

5. Repeat steps 3 and 4 for $h = 2, \ldots, p_i - 1$.

7.3.2 *Improving the Initial Estimate*

Step 2 consists of using a modified Gauss-Newton algorithm to minimize

$$\mathit{SSE}(\boldsymbol{\Gamma}_i) = \left[\mathbf{H}'_i \operatorname{vec}(\mathbf{R}_i - \widehat{\boldsymbol{\Psi}}_i)\right]' \widehat{\mathbf{M}}^{-1} \mathbf{H}'_i \operatorname{vec}(\mathbf{R}_i - \widehat{\boldsymbol{\Psi}}_i),$$

with respect to $\boldsymbol{\Gamma}_i$ subject to the constraints that $\boldsymbol{\Gamma}_i \in \mathcal{O}_{p_i}$ and $\widehat{\boldsymbol{\Psi}}_i$ is a correlation matrix, where $\widehat{\boldsymbol{\Psi}}_i = \boldsymbol{\Gamma}_i \widehat{\boldsymbol{\Lambda}}_i \boldsymbol{\Gamma}'_i$. The update of $\widehat{\boldsymbol{\Gamma}}_i$ in iteration h is $\widehat{\boldsymbol{\Gamma}}_{i,h+1} = \widehat{\boldsymbol{\Gamma}}_{i,h} \mathbf{G}(\widehat{\boldsymbol{\mu}}_{h+1})$, where

$$\begin{aligned} \widehat{\boldsymbol{\mu}}_{h+1} &= \alpha_h \left(\mathbf{X}'_{i,h} \widehat{\mathbf{M}}_i^{-1} \mathbf{X}_{i,h}\right)^{-1} \mathbf{X}'_{i,h} \widehat{\mathbf{M}}_i^{-1} \mathbf{H}'_i \operatorname{vec}(\mathbf{R}_i - \widehat{\boldsymbol{\Psi}}_{i,h}), \\ \mathbf{X}_{i,h} &= 2\mathbf{H}'_i (\widehat{\boldsymbol{\Gamma}}_{i,h} \otimes \widehat{\boldsymbol{\Gamma}}_{i,h}) \mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}, \end{aligned}$$

$\mathbf{D}^{(1)}_{\mathbf{g}_i:\boldsymbol{\mu}_i}$ is defined in Theorem 7, and $\alpha_h \in (0, 1]$. Given a value for $\widehat{\boldsymbol{\mu}}_{h+1}$, it is necessary to solve for the implicit parameters in order to compute $\mathbf{G}(\widehat{\boldsymbol{\mu}}_{h+1})$. An algorithm for solving for the implicit parameters is given in the next section.

7.3.3 *Solving for Implicit Parameters*

To reduce notational complexity, it is assumed in this section that $k = 1$ and, therefore, $p = \dot{p}$. The algorithm for obtaining an initial guess for $\boldsymbol{\Gamma}$ as well as the algorithms for minimizing the discrepancy functions require that for given values of $\boldsymbol{\mu}$ in a neighborhood of zero, $\boldsymbol{\Gamma}$, $\boldsymbol{\Lambda}$, and $\mathbf{V}$ one can solve for $\boldsymbol{\eta}$ subject to the restrictions $\mathbf{G}\mathbf{G}' = \mathbf{I}_p$ and $\mathbf{C}'_3 \operatorname{vec}(\boldsymbol{\Gamma}\mathbf{G}\boldsymbol{\Lambda}\mathbf{G}'\boldsymbol{\Gamma}' - \mathbf{I}_p) = \mathbf{0}$. See §2.3 of the manuscript for details.

The restrictions on $\boldsymbol{\eta}$ can be written as $\mathbf{k}(\boldsymbol{\eta}) = \mathbf{0}$, where

$$\begin{aligned}
\mathbf{k}(\boldsymbol{\eta}) &= \begin{pmatrix} \operatorname{vech}(\mathbf{G}\mathbf{G}' - \mathbf{I}_p) \\ \mathbf{C}_3' \operatorname{vec}(\boldsymbol{\Gamma}\mathbf{G}\boldsymbol{\Lambda}\mathbf{G}'\boldsymbol{\Gamma}' - \mathbf{I}_p) \end{pmatrix} \\
&= \mathbf{B}_0 + 2\mathbf{B}_1\boldsymbol{\eta} + \mathbf{B}_2(\boldsymbol{\eta} \otimes \boldsymbol{\eta}); \\
\mathbf{B}_0 &= \begin{pmatrix} \mathbf{D}_p' \left[(\mathbf{I}_p \otimes \operatorname{vec} \mathbf{I}_p \otimes \mathbf{I}_p)' \left(\mathbf{I}_{(p,p)} \mathbf{A}_2\mathbf{V}_2\boldsymbol{\mu} \otimes \mathbf{A}_2\mathbf{V}_2\boldsymbol{\mu}\right) - \operatorname{vec} \mathbf{I}_p\right] \\ \mathbf{C}_3' \left[(\boldsymbol{\Gamma} \otimes \operatorname{vec} \boldsymbol{\Lambda} \otimes \boldsymbol{\Gamma})' \left(\mathbf{I}_{(p,p)} \mathbf{A}_2\mathbf{V}_2\boldsymbol{\mu} \otimes \mathbf{A}_2\mathbf{V}_2\boldsymbol{\mu}\right) - \operatorname{vec} \mathbf{I}_p\right] \end{pmatrix}; \\
\mathbf{B}_1 &= \begin{pmatrix} \mathbf{D}_p' \left[(\mathbf{I}_p \otimes \operatorname{vec} \mathbf{I}_p \otimes \mathbf{I}_p)' \left(\mathbf{I}_{(p,p)} \mathbf{A}_2\mathbf{V}_2\boldsymbol{\mu} \otimes \begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\mathbf{V}_1\end{bmatrix}\right)\right] \\ \mathbf{C}_3' \left[(\boldsymbol{\Gamma} \otimes \operatorname{vec} \boldsymbol{\Lambda} \otimes \boldsymbol{\Gamma})' \left(\mathbf{I}_{(p,p)} \mathbf{A}_2\mathbf{V}_2\boldsymbol{\mu} \otimes \begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\mathbf{V}_1\end{bmatrix}\right)\right] \end{pmatrix}; \text{ and} \\
\mathbf{B}_2 &= \begin{pmatrix} \mathbf{D}_p' \left[(\mathbf{I}_p \otimes \operatorname{vec} \mathbf{I}_p \otimes \mathbf{I}_p)' \left(\mathbf{I}_{(p,p)} \begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\mathbf{V}_1\end{bmatrix} \otimes \begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\mathbf{V}_1\end{bmatrix}\right)\right] \\ \mathbf{C}_3' \left[(\boldsymbol{\Gamma} \otimes \operatorname{vec} \boldsymbol{\Lambda} \otimes \boldsymbol{\Gamma})' \left(\mathbf{I}_{(p,p)} \begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\mathbf{V}_1\end{bmatrix} \otimes \begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\mathbf{V}_1\end{bmatrix}\right)\right] \end{pmatrix};
\end{aligned}$$

where $\mathbf{D}_{\dot{p}}$ is the duplication matrix and the vech operator (Searle, 1982, §12.9) stacks the components of the columns of a matrix that are on or below the main diagonal. A modified Newton algorithm combined with an optimal line search is used to solve $\mathbf{k}(\boldsymbol{\eta}) = \mathbf{0}$. At iteration h, the value of $\boldsymbol{\eta}_h$ is updated as

$$\begin{aligned}
\boldsymbol{\eta}_{h+1} &= \boldsymbol{\eta}_h + \alpha_h\boldsymbol{\delta}_h, \text{ where} \\
\boldsymbol{\delta}_h &= -\left[\mathbf{K}(\boldsymbol{\eta}_h)\right]^{-1} \mathbf{k}(\boldsymbol{\eta}_h), \\
\mathbf{K}(\boldsymbol{\eta}_h) &= \left.\frac{\partial\, \mathbf{k}(\boldsymbol{\eta})}{\partial\, \boldsymbol{\eta}'}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta}_h} = 2\mathbf{B}_1 + 2\mathbf{B}_2(\mathbf{I}_s \otimes \boldsymbol{\eta}_h),
\end{aligned}$$

$s = (p+1)(p+2)/2 - 2$, $\mathbf{B}_2\mathbf{I}_{(\dot{p},\dot{p})} = \mathbf{B}_2$ has been used, and $\alpha_h \in (0, 1]$.

The value of α_h is chosen as the minimizer of

$$\mathit{SSE}^*(\alpha) = [\mathbf{k}(\boldsymbol{\eta}_h + \alpha\boldsymbol{\delta}_h)]'[\mathbf{k}(\boldsymbol{\eta}_h + \alpha\boldsymbol{\delta}_h)].$$

Equating the derivative of SSE^* with respect to α to zero reveals that the minimizer is one of the real solutions to the cubic equation

$$a_0 + \alpha a_1 + \alpha^2 a_2 + \alpha^3 a_3, \text{ where}$$

$$\begin{aligned}
a_0 &= -\frac{1}{2}[\mathbf{k}(\boldsymbol{\eta}_h)]'[\mathbf{k}(\boldsymbol{\eta}_h)], \\
a_1 &= \frac{1}{2}[\mathbf{k}(\boldsymbol{\eta}_h)]'[\mathbf{k}(\boldsymbol{\eta}_h)] + [\mathbf{k}(\boldsymbol{\eta}_h)]'\mathbf{B}_2(\boldsymbol{\delta}_h \otimes \boldsymbol{\delta}_h), \\
a_2 &= -\frac{3}{2}[\mathbf{k}(\boldsymbol{\eta}_h)]'\mathbf{B}_2(\boldsymbol{\delta}_h \otimes \boldsymbol{\delta}_h), \text{ and} \\
a_3 &= (\boldsymbol{\delta}_h \otimes \boldsymbol{\delta}_h)'\mathbf{B}_2'\mathbf{B}_2(\boldsymbol{\delta}_h \otimes \boldsymbol{\delta}_h).
\end{aligned}$$

The cubic equation can be written as

$$-\frac{1}{2} + \frac{1}{2}\alpha + O\left[|\mathbf{k}(\boldsymbol{\eta}_h)|\right] = \mathbf{0}.$$

Accordingly, for small $|\mathbf{k}(\boldsymbol{\eta}_h)|$, the minimizer of SSE^* will be near unity.

8 Matrix Expressions for $\mathbf{Z}_j$ and $\mathbf{K}_j$

The terms $\mathbf{Z}_j$ for $j = 1, \ldots$ in (25) are linear functions of $\mathbf{s} - \boldsymbol{\sigma} = \mathrm{vec}(\mathbf{S} - \boldsymbol{\Sigma})$ whose coefficients depend on the derivatives of $\boldsymbol{\Sigma}$ with respect to $\boldsymbol{\theta}$. For convenience, define $\ddot{\mathbf{F}}^{(1)}$, $\ddot{\mathbf{F}}^{(2)}$, $\dddot{\mathbf{F}}^{(1)}$, and $\dddot{\mathbf{F}}^{(2)}$ as

$$
\begin{aligned}
\ddot{\mathbf{F}}^{(1)} &= \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \mathbf{F}^{(1)}; \quad \ddot{\mathbf{F}}^{(2)} = \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \mathbf{F}^{(2)}; \\
\dddot{\mathbf{F}}^{(1)} &= \left(\boldsymbol{\Sigma}^{-1} \otimes \mathrm{vec}\, \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \mathbf{F}^{(1)} \text{ and } \dddot{\mathbf{F}}^{(2)} = \left(\boldsymbol{\Sigma}^{-1} \otimes \mathrm{vec}\, \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \mathbf{F}^{(2)}.
\end{aligned}
$$

Expressions for $\mathbf{Z}_j$ and $\mathbf{K}_j$ are as follows:

$$
\begin{aligned}
\mathbf{Z}_1 &= \frac{1}{2} \ddot{\mathbf{F}}^{(1)\prime} \sqrt{n}\, (\mathbf{s} - \boldsymbol{\sigma})\,; \\
\mathbf{Z}_2 &= \frac{1}{2} \mathbf{F}^{(11)\prime} \left\{\mathbf{I}_{\dot{\nu}} \otimes \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \sqrt{n}\, (\mathbf{s} - \boldsymbol{\sigma})\right\} - \dddot{\mathbf{F}}^{(1)\prime} \left\{\mathbf{F}^{(1)} \otimes \sqrt{n}(\mathbf{s} - \boldsymbol{\sigma})\right\}; \\
\mathbf{Z}_3 &= \left(\mathrm{vec}\, \ddot{\mathbf{F}}^{(1)} \otimes \ddot{\mathbf{F}}^{(1)}\right)' \left(\mathbf{I}_{\dot{p}\dot{\nu}} \otimes \mathbf{I}_{(\dot{p},\dot{p})} \otimes \mathbf{I}_{\dot{p}}\right) \left\{\mathbf{I}_{\dot{\nu}} \otimes \mathbf{F}^{(1)} \otimes \sqrt{n}(\mathbf{s} - \boldsymbol{\sigma})\right\} \\
&+ 2\, \dddot{\mathbf{F}}^{(1)\prime} \left[\mathbf{F}^{(1)} \otimes \mathbf{N}_{\dot{p}} \left(\mathbf{I}_{\dot{p}} \otimes \mathrm{vec}\, \boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_{\dot{p}}\right)' \left\{\mathbf{F}^{(1)} \otimes \sqrt{n}(\mathbf{s} - \boldsymbol{\sigma})\right\}\right] \\
&+ \frac{1}{2} \mathbf{F}^{(111)\prime} \left\{\mathbf{I}_{\dot{\nu}^2} \otimes \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \sqrt{n}\, (\mathbf{s} - \boldsymbol{\sigma})\right\} \\
&- 2\, \mathbf{F}^{(11)\prime} \left(\mathbf{I}_{\dot{\nu}} \otimes \boldsymbol{\Sigma}^{-1} \otimes \mathrm{vec}\, \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right)' \left\{\left(\mathbf{I}_{\dot{\nu}} \otimes \mathbf{F}^{(1)}\right) \mathbf{N}_{\dot{\nu}} \otimes \sqrt{n}(\mathbf{s} - \boldsymbol{\sigma})\right\} \\
&- \dddot{\mathbf{F}}^{(1)\prime} \left\{\mathbf{F}^{(2)} \otimes \sqrt{n}(\mathbf{s} - \boldsymbol{\sigma})\right\}; \\
\mathbf{K}_2 &= -\frac{1}{2} \ddot{\mathbf{F}}^{(1)\prime} \mathbf{F}^{(1)} = -\bar{\mathbf{I}}_{\boldsymbol{\theta}}; \\
\mathbf{K}_3 &= 2\, \dddot{\mathbf{F}}^{(1)\prime} \left(\mathbf{F}^{(1)} \otimes \mathbf{F}^{(1)}\right) \mathbf{N}_{\dot{\nu}} - \mathbf{F}^{(11)\prime} \left(\mathbf{I}_{\dot{\nu}} \otimes \ddot{\mathbf{F}}^{(1)}\right) \mathbf{N}_{\dot{\nu}} - \frac{1}{2} \ddot{\mathbf{F}}^{(1)\prime} \mathbf{F}^{(2)}; \text{ and} \\
\mathbf{K}_4 &= \mathbf{F}^{(11)\prime} \left\{\mathbf{I}_{\dot{\nu}} \otimes \left(\boldsymbol{\Sigma}^{-1} \otimes \mathrm{vec}\, \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right)' \left(\mathbf{F}^{(1)} \otimes \mathbf{F}^{(1)}\right)\right\} \left\{4\left(\mathbf{N}_{\dot{\nu}} \otimes \mathbf{I}_{\dot{\nu}}\right) + 2\mathbf{I}_{(\dot{\nu}^2,\dot{\nu})}\right\} \\
&+ 4\, \dddot{\mathbf{F}}^{(1)\prime} \left(\mathbf{F}^{(2)} \otimes \mathbf{F}^{(1)}\right) \left\{\left(\mathbf{I}_{\dot{\nu}} \otimes \mathbf{N}_{\dot{\nu}}\right) + \frac{1}{2} \mathbf{I}_{(\dot{\nu},\dot{\nu}^2)}\right\} \\
&- \mathbf{F}^{(1)\prime} (\boldsymbol{\Sigma}^{-1} \otimes \mathrm{vec}\, \boldsymbol{\Sigma}^{-1} \otimes \mathrm{vec}\, \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})' \left(\mathbf{F}^{(1)} \otimes \mathbf{F}^{(1)} \otimes \mathbf{F}^{(1)}\right) \\
&\times \left\{2\mathbf{I}_{(\dot{\nu},\dot{\nu}^2)} + 3\mathbf{I}_{(\dot{\nu}^2,\dot{\nu})} + 4\left(\mathbf{N}_{\dot{\nu}} \otimes \mathbf{I}_{\dot{\nu}}\right)\right\} \\
&- \mathbf{F}^{(111)\prime} \left(\mathbf{I}_{\dot{\nu}^2} \otimes \ddot{\mathbf{F}}^{(1)}\right) \left\{\left(\mathbf{I}_{\dot{\nu}} \otimes \mathbf{N}_{\dot{\nu}}\right) + \frac{1}{2} \mathbf{I}_{(\dot{\nu},\dot{\nu}^2)}\right\}
\end{aligned}
$$

$$- \quad \mathbf{F}^{(11)\prime}\left(\mathbf{I}_{\dot{\nu}} \otimes \ddot{\mathbf{F}}^{(2)}\right)\left\{(\mathbf{N}_{\dot{\nu}} \otimes \mathbf{I}_{\dot{\nu}}) + \frac{1}{2}\mathbf{I}_{(\dot{\nu}^2,\dot{\nu})}\right\} - \frac{1}{2}\,\ddot{\mathbf{F}}^{(1)\prime}\,\mathbf{F}^{(3)};$$

where $\mathbf{N}_{\dot{\nu}} = \frac{1}{2}\left(\mathbf{I}_{\dot{\nu}} + \mathbf{I}_{(\dot{\nu},\dot{\nu})}\right)$. In the current application, the expressions for $\mathbf{K}_3$ and $\mathbf{K}_4$ can be simplified because the Bartlett correction depends on these terms only through the expectations of functions such as $\mathbf{Z}_1^{*\prime}\mathbf{K}_3(\mathbf{Z}_1^* \otimes \mathbf{Z}_1^*)$ and $\mathbf{Z}_1^{*\prime}\mathbf{K}_4(\mathbf{Z}_1^* \otimes \mathbf{Z}_1^* \otimes \mathbf{Z}_1^*)$; see equation (26) in Boik (2002) for details. The simplified quantities are denoted by $\mathbf{K}_3^*$ and $\mathbf{K}_4^*$ and are defined in the next section.

9 Expectation of Likelihood Ratio Statistic

Evaluation of the expectation of the right-hand-side of equation (26) in Boik (2002) reveals that the expectation of the likelihood ratio statistic is the following:

$$\mathrm{E}\left\{D\left(\widehat{\boldsymbol{\theta}};\boldsymbol{\theta}\right)\right\} = \dim(\boldsymbol{\theta}) + \frac{1}{\bar{n}}\sum_{i=1}^{10}\zeta_i + O\left(\bar{n}^{-2}\right),$$

where

$$\begin{aligned}
\zeta_1 &= \operatorname{tr}\left\{\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime} \otimes \mathbf{N}_{\dot{p}}\right)(\mathbf{I}_{\dot{p}} \otimes \operatorname{vec}\boldsymbol{\Sigma} \otimes \mathbf{I}_{\dot{p}})\,\mathbf{F}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{U}^{(1)}\right\};\\
\mathbf{U}^{(1)} &= \mathbf{F}^{(11)\prime}\left(\mathbf{I}_{\dot{\nu}} \otimes \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) - 2\,\dddot{\mathbf{F}}^{(1)\prime}\left(\mathbf{F}^{(1)} \otimes \mathbf{I}_{\dot{p}^2}\right);\\
\zeta_2 &= \frac{1}{3}\operatorname{tr}\left\{\left(\mathbf{F}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1} \otimes \mathbf{F}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)^{\prime}\dddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{K}_3^*\right\};\\
\mathbf{K}_3^* &= 2\,\dddot{\mathbf{F}}^{(1)\prime}\left(\mathbf{F}^{(1)} \otimes \mathbf{F}^{(1)}\right) - \operatorname{dvec}\left(\ddot{\mathbf{F}}^{(1)\prime}\,\mathbf{F}^{(2)}, \dot{\nu}^2, \dot{\nu}\right)^{\prime} - \frac{1}{2}\,\ddot{\mathbf{F}}^{(1)\prime}\,\mathbf{F}^{(2)};\\
\zeta_3 &= \frac{1}{2}\operatorname{tr}\left[\left\{\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1} \otimes \mathbf{N}_{\dot{p}}\,(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})\right\}\mathbf{U}^{(1)\prime}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{U}^{(1)}\right];\\
\zeta_4 &= \mathbf{L}_1^{\prime}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{L}_1 + \mathbf{L}_2^{\prime}\left(\mathbf{I}_{(\dot{\nu},\dot{\nu})} \otimes \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\mathbf{L}_2;\\
\mathbf{L}_1 &= \frac{1}{2}\mathbf{U}^{(1)}\operatorname{vec}\left(\mathbf{F}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right); \quad \mathbf{L}_2 = \frac{1}{2}\operatorname{vec}\left\{\mathbf{U}^{(1)}\left(\mathbf{I}_{\dot{\nu}} \otimes \mathbf{F}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\right\};\\
\zeta_5 &= \frac{1}{2}\operatorname{tr}\left\{\mathbf{N}_{\dot{\nu}}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1} \otimes \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\mathbf{K}_3^{*\prime}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{K}_3^*\right\} + \frac{1}{4}\left(\operatorname{vec}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)^{\prime}\mathbf{K}_3^{*\prime}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{K}_3^*\operatorname{vec}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1};\\
\zeta_6 &= \operatorname{tr}\left\{\mathbf{N}_{\dot{\nu}}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1} \otimes \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{F}^{(1)\prime}\right)\mathbf{U}^{(1)\prime}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{K}_3^*\right\};\\
\zeta_7 &= \frac{1}{2}\left(\operatorname{vec}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)^{\prime}\mathbf{K}_3^{*\prime}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{U}^{(1)}\operatorname{vec}\left(\mathbf{F}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right);\\
\zeta_8 &= \frac{1}{3}\operatorname{tr}\left\{\mathbf{N}_{\dot{\nu}}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{F}^{(1)\prime} \otimes \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\mathbf{U}^{(2)}\right\};\\
\mathbf{U}^{(2)} &= -6\operatorname{dvec}\left(\dddot{\mathbf{F}}^{(1)\prime}, \dot{p}^2\dot{\nu}, \dot{p}^2\right)\mathbf{F}^{(2)}
\end{aligned}$$

$$
\begin{aligned}
& + && \operatorname{dvec}\left[\operatorname{dvec}\left\{\left(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\right)\mathbf{F}^{(3)}, \dot{p}^2\dot{\nu}^2, \dot{\nu}\right\}', \dot{p}^2\dot{\nu}, \dot{\nu}^2\right] \\
& + && 6(\mathbf{I}_{\dot{p}^2}\otimes\mathbf{F}^{(1)})'(\boldsymbol{\Sigma}^{-1}\otimes\operatorname{vec}\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}_{\dot{p}})\mathbf{I}_{(\dot{p},\dot{p})}\mathbf{L}_3; \\
\mathbf{L}_3 & = && \left(\boldsymbol{\Sigma}^{-1}\otimes\operatorname{vec}\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}_{\dot{p}}\right)'\left(\mathbf{F}^{(1)}\otimes\mathbf{F}^{(1)}\right); \\
\zeta_9 & = && \frac{1}{6}\left\{\operatorname{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{F}^{(1)\prime}\right)\right\}'\mathbf{U}^{(2)}\left(\operatorname{vec}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right); \\
\zeta_{10} & = && \frac{1}{6}\operatorname{tr}\left\{\mathbf{N}_{\dot{\nu}}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\mathbf{K}_4^*\right\}+\frac{1}{12}\left(\operatorname{vec}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)'\mathbf{K}_4^*\operatorname{vec}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1};\ \text{and} \\
\mathbf{K}_4^* & = && 12\left(\mathbf{F}^{(1)}\otimes\mathbf{F}^{(1)}\right)'\dddot{\mathbf{F}}^{(2)}-9\mathbf{L}_3'\mathbf{I}_{(\dot{p},\dot{p})}\mathbf{L}_3 \\
& - && 2\operatorname{dvec}\left(\ddot{\mathbf{F}}^{(1)\prime}\mathbf{F}^{(3)}, \dot{\nu}^2, \dot{\nu}^2\right)-\frac{3}{2}\mathbf{F}^{(2)\prime}\ddot{\mathbf{F}}^{(2)}.
\end{aligned}
$$

10 References

Boik, R.J. (1998). A local parameterization of orthogonal and semi-orthogonal matrices with applications. *Journal of Multivariate Analysis*, **67**, 244–276.

Boik, R.J. (2002). Spectral models for Covariance Matrices. *Biometrika*, **89**, 159–182.

Kollo, T., & Neudecker, H. (1993). Asymptotics of eigenvalues and unit-length eigenvectors of sample variance and correlation matrices. *Journal of Multivariate Analysis*, **47**, 283–300. Corrigendum, 1994, **51**, 210.

Lin, S.P., & Bendel, R.B. (1985). Algorithm AS 213. Generation of Population Correlation Matrices with Specified Eigenvalues, *Applied Statistics*, **34**, 193–198.

MacRae, E.C. (1974). Matrix derivatives with an application to an adaptive linear decision problem. *Annals of Statistics*, **2**, 337–346.

Magnus, J.R. & Neudecker, H. (1979). The commutation matrix: some properties and applications. *Annals of Statistics*, **7**, 381–394.

Magnus, J.R. & Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, rev. ed. Chichester: John Wiley & Sons.

Schott, J.R. (1997). *Matrix Analysis for Statistics*, New York: John Wiley & Sons.

Searle, S.R. (1982). *Matrix Algebra Useful for Statistics.* New York: John Wiley & Sons.

Stapleton, J.H. (1995). *Linear Statistical Models*, New York: John Wiley & Sons.