de Carvalho, M., and Davison, A. C. (2014), "Spectral Density Ratio Models for Multivariate Extremes," Journal of the American Statistical Association, 109, 764–776.

Spectral Density Ratio Models for Multivariate Extremes

(Supplementary Material)

Miguel de Carvalho and Anthony C. Davison

1 D-Dimensional Maximum Empirical Likelihood Estimator

The extension of the maximum empirical likelihood estimator of Einmahl and Segers (2009) to the D-dimensional setting is straightforward. The maximum empirical likelihood estimator $\ddot{H}(\cdot) = \sum_i \ddot{p}_i \, \delta_{\boldsymbol{v}_i}(\cdot)$ has probability masses given by the solution of

$$\begin{aligned} \max_{\boldsymbol{p} \in \mathbb{R}_+^n} \quad & \textstyle\sum_{i=1}^n \log p_i \\ \text{such that} \quad & \textstyle\sum_{i=1}^n p_i = 1, \\ & \textstyle\sum_{i=1}^n \boldsymbol{v}_i p_i = D^{-1}\mathbf{1}_D. \end{aligned} \tag{1}$$

By the method of Lagrange multipliers, the solution is given by

$$\ddot{p}_i = \frac{1}{n} \frac{1}{1 + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)}, \quad i = 1, \ldots, n,$$

where $\boldsymbol{\lambda} \in \mathbb{R}^D$ is the Lagrange multiplier associated to the marginal moment constraint in (1), defined implicitly as the solution to the equation

$$\frac{1}{n} \sum_{i=1}^n \frac{\boldsymbol{v}_i - D^{-1}\mathbf{1}_D}{1 + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)} = \mathbf{0}.$$

See Qin and Lawless (1994) for further details.

Miguel de Carvalho is Assistant Professor, Department of Statistics, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Santiago, Chile (*MdeCarvalho@mat.puc.cl*). Anthony Davison is Professor of Statistics, EPFL-FSB-MATHAA-STAT, Station 8, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (*Anthony.Davison@epfl.ch*). The research was supported by the Swiss, Portuguese and Chilean National Science Foundations. The authors thank the editor, associate editor, and referees for their careful reading and constructive comments, Vanda de Carvalho and Konstantinos Fokianos for valuable suggestions, and Jacques Ferrez for data preprocessing.

2 Additional Empirical Reports

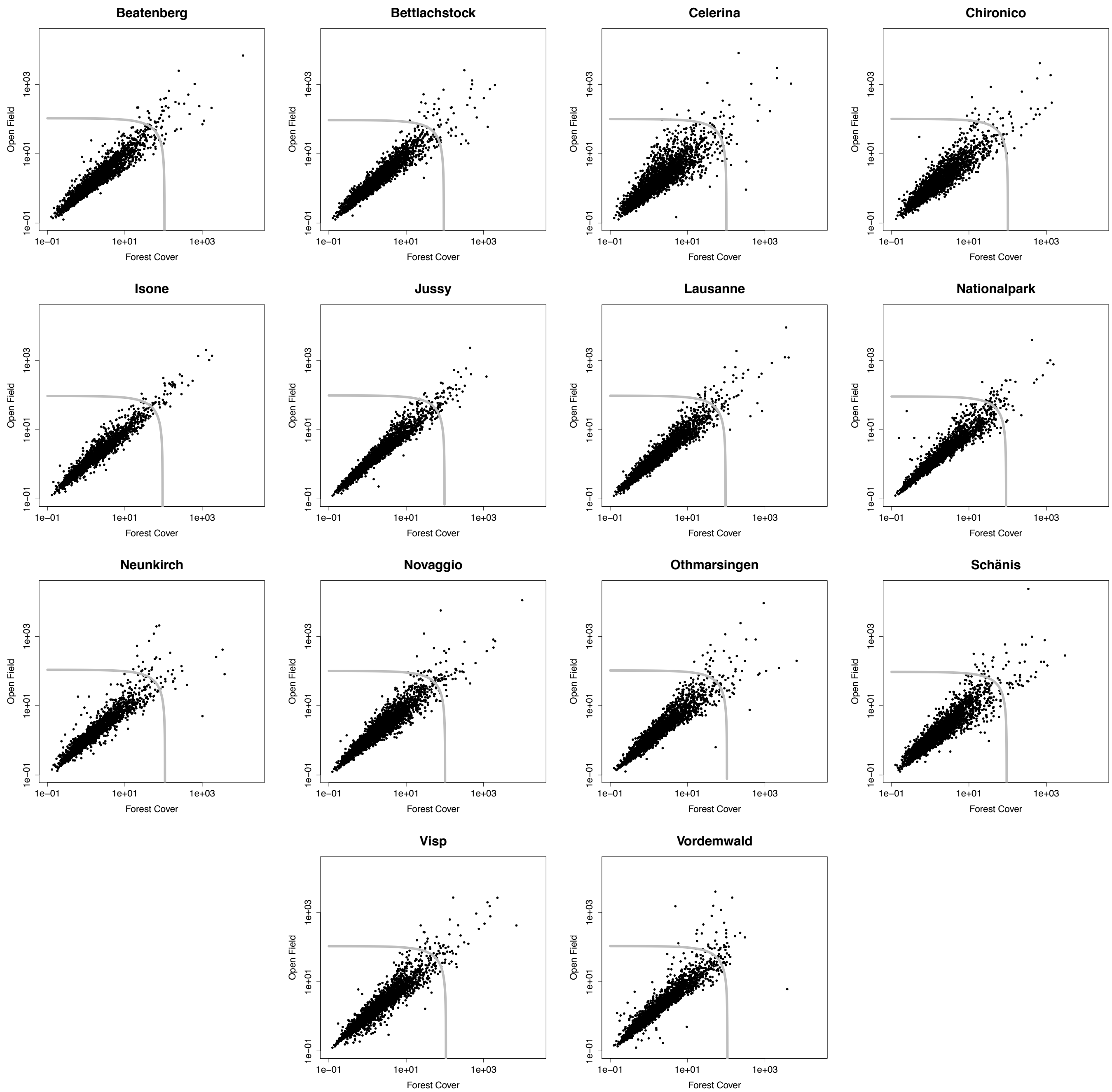

Figure 1: Scatterplot of air temperature data after transformation to unit Fréchet scale; the gray line represents the boundary threshold, at the 98% quantile, corresponding to each sample in the log-log scale, with both axes being logarithmic.

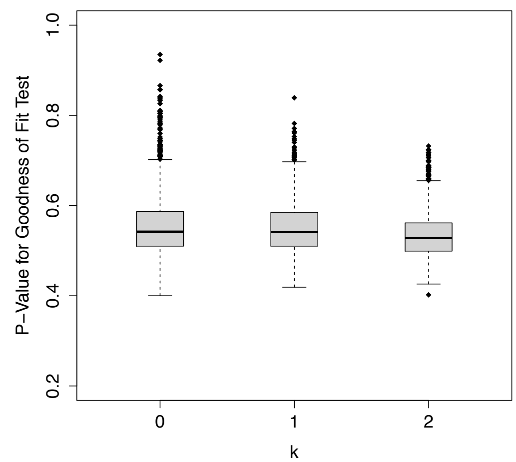

Figure 2: Boxplot of p-values for the bootstrap-based goodness of fit test in Section 2.4; each p-value is obtained by resampling 1000 times, from one simulated data set. The histogram of p-values is constructed by repeating this procedure for all the simulated data sets in Section 3.2.

3 Convex Dual Representation and the Inner Optimization Problem

The convex dual representation of the empirical likelihood problem under analysis is helpful for computational purposes. Here the necessary dual involves minimising

$$\mathcal{L}(\boldsymbol{\lambda}) = -\sum_{i=1}^{n} \log \left\{ \sum_{k=0}^{K} \exp\{\alpha_k + \beta_k c(\boldsymbol{v}_i)\}\{\rho_k + \boldsymbol{\lambda}_k^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)\} \right\},$$

subject to the linear constraints $\sum_{k=0}^{K} \exp\{\alpha_k + \beta_k c(\boldsymbol{v}_i)\}\{\rho_k + \boldsymbol{\lambda}_k^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)\} > 0$, for $i = 1, \ldots, n$. These constraints can be removed by using the pseudo-logarithmic function introduced by Owen (2001, p. 235), i.e.,

$$\log_{\#}(s) = \begin{cases} \log(s), & s > \varepsilon, \\ \log(\varepsilon) - 1.5 + 2s/\varepsilon - s^2/(2\varepsilon^2), & s \leq \varepsilon, \end{cases}$$

for some small $\varepsilon > 0$. Then the initial problem of interest simplifies into one of minimizing

$$\mathcal{L}_{\#}(\boldsymbol{\lambda}) = -\sum_{i=1}^{n} \log_{\#} \left\{ \sum_{k=0}^{K} \exp\{\alpha_k + \beta_k c(\boldsymbol{v}_i)\}\{\rho_k + \boldsymbol{\lambda}_k^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)\} \right\},$$

over $\boldsymbol{\lambda} \in \mathbb{R}^{(K+1)D}$, for which a Newton algorithm can be implemented by recursive least squares. We write the gradient and Hessian of $\mathcal{L}_{\#}$ as

$$\frac{\partial \mathcal{L}_{\#}}{\partial \boldsymbol{\lambda}} = -\boldsymbol{U}^{\mathrm{T}}\boldsymbol{y}, \qquad \frac{\partial \mathcal{L}_{\#}^2}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^{\mathrm{T}}} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{U},$$

where $\boldsymbol{U} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)^{\mathrm{T}}$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$, are defined as

$$\boldsymbol{u}_i = \left[-\log_{\#}^{(2)} \left\{ \sum_{k=0}^{K} \exp\{\alpha_k + \beta_k c(\boldsymbol{v}_i)\}\{\rho_k + \boldsymbol{\lambda}_k^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)\} \right\} \right]^{1/2} \times \quad (\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)^{\mathrm{T}},$$

$$y_i = \frac{\log_{\#}^{(1)} \left\{ \sum_{k=0}^{K} \exp\{\alpha_k + \beta_k c(\boldsymbol{v}_i)\}\{\rho_k + \boldsymbol{\lambda}_k^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)\} \right\}}{\left[-\log_{\#}^{(2)} \left\{ \sum_{k=0}^{K} \exp\{\alpha_k + \beta_k c(\boldsymbol{v}_i)\}\{\rho_k + \boldsymbol{\lambda}_k^{\mathrm{T}}(\boldsymbol{v}_i - D^{-1}\mathbf{1}_D)\} \right\} \right]^{1/2}},$$

and $\log_{\#}^{(i)}$ denotes the ith derivative of the pseudo-logarithmic function.

Numerical optimization can then be performed by updating $\boldsymbol{\lambda}$ according to the rule $\boldsymbol{\lambda} \to \boldsymbol{\lambda} + (\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U})^{-1}\boldsymbol{U}^{\mathrm{T}}y$, which uses the preceding values of the Lagrange multipliers $\boldsymbol{\lambda}$ corresponding to the marginal moment constraints, and an increment $(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U})^{-1}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{y}$; the latter is readily obtained by least squares regression of $\boldsymbol{y}$ on $\boldsymbol{U}$. See Owen (2001, sec. 3.14) for further details.

4 Auxiliary Lemmas

Lemma 1. *Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\lambda})^{\mathrm{T}}$ and $\boldsymbol{\theta}^* = (\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*, \mathbf{0})^{\mathrm{T}}$, and suppose that the conditions of Theorem 1 hold. Then*

$$-\frac{1}{n}\frac{\partial^2 \kappa}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \xrightarrow{p} \boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_{\boldsymbol{\beta\beta}} & \boldsymbol{S}_{\boldsymbol{\beta\alpha}} & \boldsymbol{S}_{\boldsymbol{\beta\lambda}} \\ & \boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}_{\boldsymbol{\alpha\lambda}} \\ & & \boldsymbol{S}_{\boldsymbol{\lambda\lambda}} \end{pmatrix}, \quad n \to \infty,$$

where the matrix $\boldsymbol{S}$ is symmetric, and, with $I(\cdot)$ the indicator function and $\rho_0 = n_0/n, \ldots, \rho_K = n_K/n$, we have

$$\begin{aligned} (\boldsymbol{S}_{\boldsymbol{\beta\beta}})_{k,m} &= I(k=m)\rho_k J_k^{cc} - \rho_k\rho_m J_{k,m}^{cc}, \quad (\boldsymbol{S}_{\boldsymbol{\beta\alpha}})_{k,m} = I(k=m)\rho_k J_k^{c} - \rho_k\rho_m J_{k,m}^{c}, \\ (\boldsymbol{S}_{\boldsymbol{\beta\lambda}})_{k,m} &= I(k=m)\rho_k J_k^{cv} - \rho_k J_{k,m}^{cv}, \quad (\boldsymbol{S}_{\boldsymbol{\alpha\alpha}})_{k,m} = I(k=m)\rho_k - \rho_k\rho_m J_{k,m}, \\ (\boldsymbol{S}_{\boldsymbol{\alpha\lambda}})_{k,m} &= -\rho_k J_{k,m}^{v}, \quad (\boldsymbol{S}_{\boldsymbol{\lambda\lambda}})_{k,m} = -J_{k,m}^{vv}, \end{aligned}$$

for $k,m=1,\ldots,K$ when considering $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $k,m=0,\ldots,K$ when considering $\boldsymbol{\lambda}$, where

$$\begin{aligned}
J_k^{cc} &= \int_{S_D} c^2(\boldsymbol{v})\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\mathrm{d}H_0(\boldsymbol{v}),\\
J_{k,m}^{cc} &= \int_{S_D} c^2(\boldsymbol{v})\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{v})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v})\}}\mathrm{d}H_0(\boldsymbol{v}),\\
J_k^{c} &= \int_{S_D} c(\boldsymbol{v})\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\mathrm{d}H_0(\boldsymbol{v}),\\
J_{k,m}^{c} &= \int_{S_D} c(\boldsymbol{v})\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{v})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v})\}}\mathrm{d}H_0(\boldsymbol{v}),\\
J_k^{cv} &= \int_{S_D} c(\boldsymbol{v})(\boldsymbol{v}-D^{-1}\mathbf{1}_D)\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\mathrm{d}H_0(\boldsymbol{v}),\\
J_{k,m}^{cv} &= \int_{S_D} c(\boldsymbol{v})(\boldsymbol{v}-D^{-1}\mathbf{1}_D)\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{v})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v})\}}\mathrm{d}H_0(\boldsymbol{v}),\\
J_{k,m} &= \int_{S_D} \frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{v})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v})\}}\mathrm{d}H_0(\boldsymbol{v}),\\
J_{k,m}^{v} &= \int_{S_D} (\boldsymbol{v}-D^{-1}\mathbf{1}_D)\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{v})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v})\}}\mathrm{d}H_0(\boldsymbol{v}),\\
J_{k,m}^{vv} &= \int_{S_D} (\boldsymbol{v}-D^{-1}\mathbf{1}_D)(\boldsymbol{v}-D^{-1}\mathbf{1}_D)^{\mathrm{T}}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{v})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{v})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v})\}}\mathrm{d}H_0(\boldsymbol{v}).
\end{aligned}$$

These expressions are understood to be evaluated at $\alpha_k=\alpha_k^*$, $\beta_k=\beta_k^*$, for $k,m=1,\ldots,K$.

Proof. For compactness let $Q_i=\sum_{l=0}^{K}\exp\{\alpha_l+\beta_l c(\boldsymbol{v}_i)\}\{\rho_l+\boldsymbol{\lambda}_l^{\mathrm{T}}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)\}$, $c_i=c(\boldsymbol{v}_i)$, $c_{k,j}=c(\boldsymbol{w}_{k,j})$ and $d_{k,i}=\rho_k+\boldsymbol{\lambda}_k^{\mathrm{T}}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)$, and note that when $(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\lambda})$ equals $(\boldsymbol{\beta}^*,\boldsymbol{\alpha}^*,\mathbf{0})$, we have $d_{k,i}=\rho_k$ and $Q_i=\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v}_i)\}$. Definition (22) in the paper of the auxiliary function κ implies that

$$\begin{aligned}
-\frac{\partial\kappa}{\partial\alpha_k} &= \sum_{i=1}^{n} e^{\alpha_k+\beta_k c_i}d_{k,i}/Q_i-n_k, &(2)\\
-\frac{\partial\kappa}{\partial\beta_k} &= \sum_{i=1}^{n} c_i e^{\alpha_k+\beta_k c_i}d_{k,i}/Q_i-\sum_{j=1}^{n_k}c_{k,j}, &(3)\\
-\frac{\partial\kappa}{\partial\boldsymbol{\lambda}_k} &= \sum_{i=1}^{n}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)e^{\alpha_k+\beta_k c_i}/Q_i. &(4)
\end{aligned}$$

If these expressions are evaluated at $(\boldsymbol{\beta}^*,\boldsymbol{\alpha}^*,\mathbf{0})$, then since $\sum_{i=1}^{n}(\cdot)\equiv\sum_{l=0}^{K}\sum_{j=1}^{n_l}(\cdot)$ and $n_l=n\rho_l$, the expectation of

(2) equals

$$
\begin{aligned}
\mathrm{E}\left[\sum_{i=1}^{n}\frac{\rho_k\exp\{\alpha_k+\beta_k c(\boldsymbol{v}_i)\}}{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v}_i)\}}\right] &= \rho_k\int_{S_D}\sum_{l=0}^{K}\sum_{j=1}^{n_l}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{w}_{l,j})\}}{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w}_{l,j})\}}\mathrm{d}H_l(\boldsymbol{w}_{l,j})-n_k \\
&= \rho_k\int_{S_D}\sum_{l=0}^{K}\frac{n_l\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}}{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}\mathrm{d}H_0(\boldsymbol{w})-n_k \\
&= \rho_k\int_{S_D}\sum_{l=0}^{K}\frac{n_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}\mathrm{d}H_k(\boldsymbol{w})-n_k \\
&= n\rho_k\int_{S_D}\frac{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}\mathrm{d}H_k(\boldsymbol{w})-n_k \\
&= n\rho_k-n_k \\
&= 0. \qquad (5)
\end{aligned}
$$

Similar computations show that (3) and (4) also have zero expectations at $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*, \mathbf{0})$; the computation for (4) uses the moment constraint (3) of the paper.

The second derivatives may be written as

$$
\begin{aligned}
-\frac{\partial^2\kappa}{\partial\alpha_k\partial\alpha_m} &= I(k=m)\sum_{i=1}^{n}e^{\alpha_k+\beta_k c_i}d_{k,i}/Q_i-\sum_{i=1}^{n}e^{\alpha_k+\beta_k c_i}e^{\alpha_m+\beta_m c_i}d_{k,i}d_{m,i}/Q_i^2, \\
-\frac{\partial^2\kappa}{\partial\beta_k\partial\alpha_m} &= I(k=m)\sum_{i=1}^{n}c_i e^{\alpha_k+\beta_k c_i}d_{k,i}/Q_i-\sum_{i=1}^{n}c_i e^{\alpha_k+\beta_k c_i}e^{\alpha_m+\beta_m c_i}d_{k,i}d_{m,i}/Q_i^2, \\
-\frac{\partial^2\kappa}{\partial\boldsymbol{\lambda}_k\partial\alpha_m} &= I(k=m)\sum_{i=1}^{n}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)e^{\alpha_k+\beta_k c_i}/Q_i-\sum_{i=1}^{n}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)d_{m,i}e^{\alpha_k+\beta_k c_i}e^{\alpha_m+\beta_m c_i}/Q_i^2 \\
-\frac{\partial^2\kappa}{\partial\beta_k\partial\beta_m} &= I(k=m)\sum_{i=1}^{n}c_i^2 e^{\alpha_k+\beta_k c_i}d_{k,i}/Q_i-\sum_{i=1}^{n}c_i^2 e^{\alpha_k+\beta_k c_i}e^{\alpha_m+\beta_m c_i}d_{k,i}d_{m,i}/Q_i^2, \\
-\frac{\partial^2\kappa}{\partial\boldsymbol{\lambda}_k\partial\beta_m} &= I(k=m)\sum_{i=1}^{n}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)c_i e^{\alpha_k+\beta_k c_i}/Q_i-\sum_{i=1}^{n}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)c_i d_{m,i}e^{\alpha_k+\beta_k c_i}e^{\alpha_m+\beta_m c_i}/Q_i^2, \\
-\frac{\partial^2\kappa}{\partial\boldsymbol{\lambda}_k\partial\boldsymbol{\lambda}_m^{\mathrm{T}}} &= -\sum_{i=1}^{n}(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)(\boldsymbol{v}_i-D^{-1}\mathbf{1}_D)^{\mathrm{T}}e^{\alpha_k+\beta_k c_i}e^{\alpha_m+\beta_m c_i}/Q_i^2.
\end{aligned}
$$

The weak law of large numbers implies that each of these sums of independent variates, when renormalised by division by n, will converge to the corresponding expectation, if this is finite. When the expressions are evaluated at $(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\lambda}) = (\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*, \mathbf{0})$, we find that the manipulations like those leading to (5) give, for example,

$$
\mathrm{E}\left(-\frac{1}{n}\frac{\partial^2\kappa}{\partial\alpha_k\partial\alpha_m}\right) = (\boldsymbol{S}_{\boldsymbol{\alpha}\boldsymbol{\alpha}})_{k,m} = I(k=m)\rho_k-\rho_k\rho_m J_{k,m},
$$

where the value of the first term stems from the computation leading to (5) and the second appears because

$$\begin{aligned}
J_{k,m} &= \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{v}_i)\}\exp\{\alpha_m+\beta_m c(\boldsymbol{v}_i)\}}{\left[\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{v}_i)\}\right]^2}\right]\\
&= \frac{1}{n}\int_{S_D}\sum_{l=0}^{K}\sum_{j=1}^{n_l}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{w}_{l,j})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{w}_{l,j})\}}{\left[\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w}_{l,j})\}\right]^2}\mathrm{d}H_l(\boldsymbol{w}_{l,j})\\
&= \frac{1}{n}\int_{S_D}\sum_{l=0}^{K}\frac{n_l\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{w})\}}{\left[\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}\right]^2}\mathrm{d}H_l(\boldsymbol{w})\\
&= \int_{S_D}\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{w})\}\frac{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}{\left[\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}\right]^2}\mathrm{d}H_0(\boldsymbol{w})\\
&= \int_{S_D}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{w})\}}{\sum_{l=0}^{K}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}\mathrm{d}H_0(\boldsymbol{w}).
\end{aligned}\tag{6}$$

The other expectations are computed similarly, with the first term vanishing from the computation for $(\boldsymbol{S}_{\boldsymbol{\alpha\lambda}})_{k,m}$ because the expected value of (4) equals zero. □

Lemma 2. *Let $\boldsymbol{\theta}=(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\lambda})^{\mathrm{T}}$ and $\boldsymbol{\theta}^*=(\boldsymbol{\beta}^*,\boldsymbol{\alpha}^*,\boldsymbol{0})^{\mathrm{T}}$, and suppose that the conditions of Theorem 1 hold. Then*

$$\frac{1}{\sqrt{n}}\frac{\partial\kappa}{\partial\boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}\xrightarrow{d}N(\boldsymbol{0},\boldsymbol{V}),\quad \boldsymbol{V}=\boldsymbol{S}-\boldsymbol{T}_0-\boldsymbol{T}_1-\boldsymbol{T}_2,\quad n\to\infty,\tag{7}$$

where $\boldsymbol{S}$ is defined in Lemma 1, and, with $\boldsymbol{R}=diag(\rho_1,\ldots,\rho_K)$,

$$\boldsymbol{T}_0=\begin{pmatrix}\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{S}_{\boldsymbol{\beta\lambda}}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\\ \boldsymbol{S}_{\boldsymbol{\lambda\beta}}&\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}&2\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}\end{pmatrix},\quad \boldsymbol{T}_1=\boldsymbol{S}\begin{pmatrix}\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{R}^{-1}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\end{pmatrix}\boldsymbol{S},\quad \boldsymbol{T}_2=\rho_0^{-1}\boldsymbol{S}\begin{pmatrix}\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\end{pmatrix}\boldsymbol{S}.$$

Proof. The vector random variable

$$\frac{1}{\sqrt{n}}\frac{\partial\kappa}{\partial\boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}$$

is a sum of independent terms, and results in Lemma 1 imply that its components (2), (3) and (4) have expectations zero. Thus provided its variance matrix is finite, the result follows by the central limit theorem. We thus must show that its variance matrix is of form (7). The first step in establishing this is the computation of the variances and covariances of (2), (3) and (4) when $\boldsymbol{\theta} = \boldsymbol{\theta}^*$. At that point, and in the notation used in the proof of Lemma 1, we have

$$\begin{aligned}
\frac{1}{n}\mathrm{cov}\left(\frac{\partial\kappa}{\partial\alpha_k}, \frac{\partial\kappa}{\partial\alpha_m}\right) &= \frac{1}{n}\mathrm{cov}\left\{\sum_{i=1}^{n} e^{\alpha_k+\beta_k c(\boldsymbol{v}_i)}\rho_k/Q_i, \sum_{i=1}^{n} e^{\alpha_m+\beta_m c(\boldsymbol{v}_i)}\rho_m/Q_i\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{cov}\left\{e^{\alpha_k+\beta_k c(\boldsymbol{v}_i)}\rho_k/Q_i, e^{\alpha_m+\beta_m c(\boldsymbol{v}_i)}\rho_m/Q_i\right\} \\
&= \frac{1}{n}\sum_{l=0}^{K}\sum_{j=1}^{n_l}\mathrm{cov}\left\{e^{\alpha_k+\beta_k c(\boldsymbol{w}_{l,j})}\rho_k/Q_i, e^{\alpha_m+\beta_m c(\boldsymbol{w}_{l,j})}\rho_m/Q_i\right\} \\
&= \rho_k\rho_m\sum_{l=0}^{K}\rho_l\Bigg\{\int_{S_D}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{w})\}}{\left[\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}\right]^2}\mathrm{d}H_l(\boldsymbol{w}) \\
&\quad -\int_{S_D}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}\mathrm{d}H_l(\boldsymbol{w})\int_{S_D}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}\exp\{\alpha_m+\beta_m c(\boldsymbol{w})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}\mathrm{d}H_l(\boldsymbol{w})\Bigg\} \\
&= \rho_k\rho_m\left(J_{k,m} - \sum_{l=0}^{K}\rho_l J_{k,l}J_{l,m}\right); \qquad (8)
\end{aligned}$$

the equalities successively following from the definition of covariance, the independence of the observations, re-expression in terms of the $K+1$ samples, the definition of covariance and (6). Now

$$\sum_{l=0}^{K}\rho_l J_{l,k} = \sum_{l=0}^{K}\rho_l\int_{S_D}\frac{\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}}{\sum_{l=0}^{K-1}\rho_l\exp\{\alpha_l+\beta_l c(\boldsymbol{w})\}}\mathrm{d}H_l(\boldsymbol{w}) = \int_{S_D}\exp\{\alpha_k+\beta_k c(\boldsymbol{w})\}\mathrm{d}H_0(\boldsymbol{w}) = 1,$$

and this and similar computations yield

$$\rho_0 J_{k,0} = 1 - \sum_{l=1}^{K}\rho_l J_{k,l}, \quad \rho_0 J_{k,0}^v = -\sum_{l=1}^{K}\rho_l J_{k,l}^v, \quad \rho_0 J_{k,0}^c = J_k^c - \sum_{l=1}^{K}\rho_l J_{k,l}^c. \qquad (9)$$

It follows that the bracketed term in (8) may be written

$$J_{k,m} - \sum_{l=1}^{K}\rho_l J_{k,l}J_{l,m} - \rho_0^{-1}\left(1-\sum_{l=1}^{K}\rho_l J_{k,l}\right)\left(1-\sum_{l=1}^{K}\rho_l J_{l,m}\right)$$

and hence (8) may be expressed in matrix terms as

$$\boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \boldsymbol{RJR} - \boldsymbol{RJRJR} - \rho_0^{-1}\boldsymbol{R}(\boldsymbol{I}-\boldsymbol{JR})\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{RJ})\boldsymbol{R},$$

where $\boldsymbol{J}$ is the $K\times K$ matrix with (k,m) element $J_{k,m}$, for $k,m = 1,\ldots,K$.

Note that

$$\boldsymbol{T}_1 = \boldsymbol{S}\begin{pmatrix}\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}^{-1} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{S} = \begin{pmatrix}\boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\beta}} & \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\\ \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\beta}} & \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\\ \boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\beta}} & \boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\end{pmatrix}$$

and

$$\boldsymbol{T}_2 = \rho_0^{-1}\boldsymbol{S}\begin{pmatrix}\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{1}\boldsymbol{1}^{\mathrm{T}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{S} = \rho_0^{-1}\begin{pmatrix}\boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\beta}} & \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\\ \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\beta}} & \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\\ \boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\beta}} & \boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\end{pmatrix}.$$

Since $\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} = \boldsymbol{R} - \boldsymbol{RJR}$, the $(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ part of (7) equals

$$\begin{aligned} \boldsymbol{S}_{\boldsymbol{\alpha\alpha}} - \boldsymbol{0} - \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{R}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} - \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} &= \boldsymbol{R} - \boldsymbol{RJR} - (\boldsymbol{R} - \boldsymbol{RJR})\boldsymbol{R}^{-1}(\boldsymbol{R} - \boldsymbol{RJR}) - \rho_0^{-1}(\boldsymbol{R} - \boldsymbol{RJR})\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}(\boldsymbol{R} - \boldsymbol{RJR}) \\ &= \boldsymbol{R} - \boldsymbol{RJR} - \boldsymbol{R} + 2\boldsymbol{RJR} - \boldsymbol{RJRJR} - \rho_0^{-1}\boldsymbol{R}(\boldsymbol{I} - \boldsymbol{JR})\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{RJ})\boldsymbol{R} \\ &= \boldsymbol{V}_{\boldsymbol{\alpha\alpha}}. \end{aligned}$$

Tedious computations along the same lines establish that, in similar notation and with a subscript d indicating a diagonal matrix (so, for example, $\boldsymbol{J}_d^{cc}$ is the diagonal of the $K \times K$ matrix $\boldsymbol{J}^{cc}$, whose (k, m) element is $J_{k,m}^{cc}$), we have

$$\begin{aligned} \boldsymbol{S}_{\boldsymbol{\beta\beta}} &= \boldsymbol{RJ}_d^{cc} - \boldsymbol{RJ}^{cc}\boldsymbol{R}, \\ \boldsymbol{V}_{\boldsymbol{\beta\beta}} &= \boldsymbol{RJ}_d^{cc} - \boldsymbol{RJ}^{cc}\boldsymbol{R} + \boldsymbol{RJ}^{c}\boldsymbol{J}_d^{c}\boldsymbol{R} + \boldsymbol{RJ}_d^{c}\boldsymbol{J}_d^{cc}\boldsymbol{R} - \boldsymbol{R}(\boldsymbol{J}_d^{c})^2 - \boldsymbol{RJ}^{c}\boldsymbol{RJ}^{c}\boldsymbol{R} - \rho_0^{-1}\boldsymbol{R}(\boldsymbol{J}_d^{c} - \boldsymbol{J}^{c}\boldsymbol{R})\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}(\boldsymbol{J}_d^{c} - \boldsymbol{RJ}^{c})\boldsymbol{R}, \\ \boldsymbol{S}_{\boldsymbol{\beta\alpha}} &= \boldsymbol{RJ}_d^{c} - \boldsymbol{RJ}^{c}\boldsymbol{R}, \quad \boldsymbol{V}_{\boldsymbol{\beta\alpha}} = \boldsymbol{RJ}_d^{c}\boldsymbol{JR} - \boldsymbol{RJ}^{c}\boldsymbol{RJR} - \rho_0^{-1}\boldsymbol{R}(\boldsymbol{J}_d^{c} - \boldsymbol{J}^{c}\boldsymbol{R})\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{RJ})\boldsymbol{R}, \\ \boldsymbol{S}_{\boldsymbol{\beta\lambda}} &= \boldsymbol{RJ}_d^{cv} - \boldsymbol{RJ}^{cv}, \quad \boldsymbol{V}_{\boldsymbol{\beta\lambda}} = \boldsymbol{RJ}_d^{c}\boldsymbol{J}^{v} - \boldsymbol{RJ}^{c}\boldsymbol{RJ}^{v} - \rho_0^{-1}\boldsymbol{R}(\boldsymbol{J}_d^{c} - \boldsymbol{J}^{c}\boldsymbol{R})\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}(-\boldsymbol{RJ}^{v}), \\ \boldsymbol{S}_{\boldsymbol{\alpha\lambda}} &= -\boldsymbol{RJ}^{v}, \quad \boldsymbol{V}_{\boldsymbol{\alpha\lambda}} = \boldsymbol{RJ}^{v} - \boldsymbol{RJRJ}^{v} + \rho_0^{-1}\boldsymbol{R}(\boldsymbol{I} - \boldsymbol{JR})\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{RJ}^{v}, \\ \boldsymbol{S}_{\boldsymbol{\lambda\lambda}} &= -\boldsymbol{J}^{vv}, \quad \boldsymbol{V}_{\boldsymbol{\lambda\lambda}} = \boldsymbol{J}^{vv} - \boldsymbol{J}^{v}\boldsymbol{RJ}^{v} - \rho_0^{-1}\boldsymbol{J}^{v}\boldsymbol{R}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\boldsymbol{RJ}^{v}, \end{aligned}$$

from which the stated relation $\boldsymbol{V} = \boldsymbol{S} - \boldsymbol{T}_0 - \boldsymbol{T}_1 - \boldsymbol{T}_2$ follows after some matrix algebra. Since all the required matrices are finite by hypothesis, the result follows. □

The decomposition of $\boldsymbol{V}$ in Lemma 2 differs from that of Huang and Rathouz (2012, Lemma 3) due to an unfortunate error, but a decomposition similar to eq. (7) also holds in their setting.

Lemma 3. *Let $\boldsymbol{S}$ be as in Lemma 1, but partitioned according to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma} = (\boldsymbol{\alpha}, \boldsymbol{\lambda})^{\mathrm{T}}$ as*

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_{\boldsymbol{\beta\beta}} & \boldsymbol{S}_{\boldsymbol{\beta\gamma}} \\ \boldsymbol{S}_{\boldsymbol{\gamma\beta}} & \boldsymbol{S}_{\boldsymbol{\gamma\gamma}} \end{pmatrix}.$$

Then, provided the necessary inverses exist,

$$(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{S}(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} = \boldsymbol{S}_{\boldsymbol{\beta\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}, \tag{10}$$

$$(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{T}_0(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} = \boldsymbol{0}, \tag{11}$$

$$(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{T}_1(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} = \boldsymbol{0}, \tag{12}$$

$$(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{T}_2(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} = \boldsymbol{0}. \tag{13}$$

Proof. To establish (10), we write

$$\begin{aligned} (\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{S}(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} &= (\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}) \begin{pmatrix} \boldsymbol{S}_{\boldsymbol{\beta\beta}} & \boldsymbol{S}_{\boldsymbol{\beta\gamma}} \\ \boldsymbol{S}_{\boldsymbol{\gamma\beta}} & \boldsymbol{S}_{\boldsymbol{\gamma\gamma}} \end{pmatrix} (\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} \\ &= (\boldsymbol{S}_{\boldsymbol{\beta\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}, \boldsymbol{0})(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} \\ &= \boldsymbol{S}_{\boldsymbol{\beta\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}. \end{aligned}$$

For (11), we write

$$(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{T}_0(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} = (\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\begin{pmatrix} \boldsymbol{0} & \boldsymbol{S}_{\boldsymbol{\beta\gamma}}^0 \\ \boldsymbol{S}_{\boldsymbol{\gamma\beta}}^0 & \boldsymbol{S}_{\gamma\gamma} + \boldsymbol{A} \end{pmatrix}(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}}$$

where

$$\boldsymbol{S}_{\boldsymbol{\beta\gamma}}^0 = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{S}_{\boldsymbol{\beta\lambda}} \end{pmatrix}, \quad \boldsymbol{S}_{\boldsymbol{\gamma\beta}}^0 = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{S}_{\boldsymbol{\lambda\beta}} \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} -\boldsymbol{S}_{\boldsymbol{\alpha\alpha}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_{\boldsymbol{\lambda\lambda}} \end{pmatrix}.$$

Some matrix algebra yields

$$(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{T}_0(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} = -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}^0 - \boldsymbol{S}_{\boldsymbol{\beta\gamma}}^0\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}} + \boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}} + \boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{A}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}. \tag{14}$$

The usual expression for the inverse of a partitioned matrix (Mardia et al., 1979, p. 459) gives

$$\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1} = \begin{pmatrix} \boldsymbol{S}^{\boldsymbol{\alpha\alpha}} & -\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1} \\ -\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}} & \boldsymbol{S}^{\boldsymbol{\lambda\lambda}} \end{pmatrix} \tag{15}$$

where

$$(\boldsymbol{S}^{\boldsymbol{\alpha\alpha}})^{-1} = \boldsymbol{S}_{\boldsymbol{\alpha\alpha}} - \boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}, \quad (\boldsymbol{S}^{\boldsymbol{\lambda\lambda}})^{-1} = \boldsymbol{S}_{\boldsymbol{\lambda\lambda}} - \boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}, \quad \boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1} = \boldsymbol{S}_{\boldsymbol{\alpha\alpha}}^{-1}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}^{\boldsymbol{\lambda\lambda}}. \tag{16}$$

Inserting (15) into the expressions in (14) yields

$$\begin{aligned}
\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}^0 &= \boldsymbol{S}_{\beta\boldsymbol{\lambda}}\boldsymbol{S}^{\boldsymbol{\lambda\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\beta}}, \\
\boldsymbol{S}_{\boldsymbol{\beta\gamma}}^0\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}} &= \boldsymbol{S}_{\beta\boldsymbol{\lambda}}\boldsymbol{S}^{\boldsymbol{\lambda\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\beta}}, \\
\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}} &= \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\beta}} + \boldsymbol{S}_{\boldsymbol{\beta\lambda}}\boldsymbol{S}^{\boldsymbol{\lambda\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\beta}}, \\
\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{A}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}} &= \boldsymbol{S}_{\boldsymbol{\beta\alpha}}(\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}} - \boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}})\boldsymbol{S}_{\boldsymbol{\alpha\beta}} \\
&\quad + \boldsymbol{S}_{\boldsymbol{\beta\lambda}}(\boldsymbol{S}^{\boldsymbol{\lambda\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}\boldsymbol{S}^{\boldsymbol{\lambda\lambda}} - \boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1})\boldsymbol{S}_{\boldsymbol{\lambda\beta}} \\
&\quad - \boldsymbol{S}_{\boldsymbol{\beta\alpha}}(\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}^{\boldsymbol{\lambda\lambda}} - \boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1})\boldsymbol{S}_{\boldsymbol{\lambda\beta}} \\
&\quad - \boldsymbol{S}_{\boldsymbol{\beta\lambda}}(\boldsymbol{S}^{\boldsymbol{\lambda\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}} - \boldsymbol{S}_{\boldsymbol{\lambda\lambda}}^{-1}\boldsymbol{S}_{\boldsymbol{\lambda\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}})\boldsymbol{S}_{\boldsymbol{\alpha\beta}} \\
&= \boldsymbol{S}_{\boldsymbol{\beta\lambda}}\boldsymbol{S}^{\boldsymbol{\lambda\lambda}}\boldsymbol{S}_{\boldsymbol{\lambda\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\alpha}}\boldsymbol{S}^{\boldsymbol{\alpha\alpha}}\boldsymbol{S}_{\boldsymbol{\alpha\beta}},
\end{aligned}$$

where the final equality follows from using (16). On inserting these expressions into the right-hand side of (14), we obtain $\mathbf{0}$, as required to establish (11).

For (12), we have

$$\begin{aligned}
(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{T}_1(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} &= (\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1})\boldsymbol{S}\begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q} \end{pmatrix}\boldsymbol{S}(\boldsymbol{I}, -\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}})^{\mathrm{T}} \\
&= (\boldsymbol{S}_{\boldsymbol{\beta\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}, \mathbf{0})\begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q} \end{pmatrix}(\boldsymbol{S}_{\boldsymbol{\beta\beta}} - \boldsymbol{S}_{\boldsymbol{\beta\gamma}}\boldsymbol{S}_{\boldsymbol{\gamma\gamma}}^{-1}\boldsymbol{S}_{\boldsymbol{\gamma\beta}}, \mathbf{0})^{\mathrm{T}} \\
&= \mathbf{0},
\end{aligned}$$

where $\boldsymbol{Q}$ is the lower right part of the central matrix in $\boldsymbol{T}_1$, when the latter is partitioned according to $(\boldsymbol{\beta}, \boldsymbol{\gamma})$. Clearly the same argument applies if $\boldsymbol{T}_1$ is replaced by $\boldsymbol{T}_2$, and this yields (13). □

5 Proof of Corollary 2

Our argument is the same as one of Huang and Rathouz, and is given here for ease of reference.

Proof. Let $\boldsymbol{\beta}^* = (\boldsymbol{\beta}_1^0, \boldsymbol{\beta}_2) = (\mathbf{0}, \boldsymbol{\beta}_2) \in \mathbb{B}$ denote the value of $\boldsymbol{\beta}$ giving rise to the data, and note that if $\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{\mathbb{B}} = O_{\mathrm{p}}(n^{-1/2})$, which we shall establish below, then a Taylor series expansion of $\ell_{\mathrm{p}}(\hat{\boldsymbol{\beta}}_{\mathbb{B}})$ around $\hat{\boldsymbol{\beta}}$ yields

$$\begin{aligned} 2\{\ell_{\mathrm{p}}(\hat{\boldsymbol{\beta}}) - \ell_{\mathrm{p}}(\hat{\boldsymbol{\beta}}_{\mathbb{B}})\} &= \sqrt{n}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{\mathbb{B}})^{\mathrm{T}} \left(-\frac{1}{n} \frac{\mathrm{d}^2 \ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta} \mathrm{d}\boldsymbol{\beta}^{\mathrm{T}}} \right)\bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} \sqrt{n}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{\mathbb{B}}) + o_{\mathrm{p}}(1) \\ &= \sqrt{n}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{\mathbb{B}})^{\mathrm{T}}\ \boldsymbol{\Sigma}\ \sqrt{n}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{\mathbb{B}}) + o_{\mathrm{p}}(1), \end{aligned} \tag{17}$$

since $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^* = O_{\mathrm{p}}(n^{-1/2})$.

We now partition the $K \times K$ matrix $\boldsymbol{\Sigma}$ conformally with $\boldsymbol{\beta}^*$, with $\boldsymbol{\Sigma}_{11}$ an $m \times m$ matrix, and write

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \quad \boldsymbol{H} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}.$$

The last $K - m$ elements of $\mathrm{d}\ell_{\mathrm{p}}/\mathrm{d}\boldsymbol{\beta}|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathbb{B}}}$ equal zero, so $\boldsymbol{H}\ \mathrm{d}\ell_{\mathrm{p}}/\mathrm{d}\boldsymbol{\beta}|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathbb{B}}} = \mathbf{0}$.

The argument to establish part 1) of Theorem 1, but applied within the submodel $\mathbb{B}$, yield

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{\mathbb{B}} - \boldsymbol{\beta}^*) = \boldsymbol{H} \left(\frac{1}{\sqrt{n}} \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} \right) + o_{\mathrm{p}}(1), \tag{18}$$

since the first m elements of both $\hat{\boldsymbol{\beta}}_{\mathbb{B}}$ and $\boldsymbol{\beta}^*$ equal zero. The term in parentheses on the right of (18) has a limiting Gaussian distribution with mean zero, by the same argument. Thus

$$\begin{aligned} \frac{1}{\sqrt{n}} \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathbb{B}}} &= \frac{1}{\sqrt{n}} \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} + \left(\frac{1}{n} \frac{\mathrm{d}^2 \ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta} \mathrm{d}\boldsymbol{\beta}^{\mathrm{T}}}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} \right) \sqrt{n}(\hat{\boldsymbol{\beta}}_{\mathbb{B}} - \boldsymbol{\beta}^*) + o_{\mathrm{p}}(1) \\ &= \frac{1}{\sqrt{n}} \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} - \boldsymbol{\Sigma}\ \sqrt{n}(\hat{\boldsymbol{\beta}}_{\mathbb{B}} - \boldsymbol{\beta}^*) + o_{\mathrm{p}}(1) \\ &= (\boldsymbol{I} - \boldsymbol{\Sigma}\boldsymbol{H}) \left(\frac{1}{\sqrt{n}} \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} \right) + o_{\mathrm{p}}(1) \\ &\stackrel{\mathrm{d}}{\to} (\boldsymbol{I} - \boldsymbol{\Sigma}\boldsymbol{H})\boldsymbol{Z}, \quad n \to \infty, \end{aligned} \tag{19}$$

where we have used the consistency of the rescaled profile information matrix and (18), and $\boldsymbol{Z} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$.

Now both $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}_{\mathbb{B}}$ are $\sqrt{n}$-consistent for $\boldsymbol{\beta}^*$, so $\hat{\boldsymbol{\beta}}_{\mathbb{B}} - \hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_{\mathbb{B}} - \boldsymbol{\beta}^*) - (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) = O_{\mathrm{p}}(n^{-1/2})$, as stated above. This yields that

$$\begin{aligned} \frac{1}{\sqrt{n}} \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathbb{B}}} &= \frac{1}{\sqrt{n}} \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} + \left(\frac{1}{n} \frac{\mathrm{d}^2 \ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta} \mathrm{d}\boldsymbol{\beta}^{\mathrm{T}}}\bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} \right) \sqrt{n}(\hat{\boldsymbol{\beta}}_{\mathbb{B}} - \hat{\boldsymbol{\beta}}) + o_{\mathrm{p}}(1) \\ &= \boldsymbol{\Sigma}\ \sqrt{n}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{\mathbb{B}}) + o_{\mathrm{p}}(1), \end{aligned}$$

and this yields

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{\mathbb{B}}) = \boldsymbol{\Sigma}^{-1} \, \frac{1}{\sqrt{n}} \, \frac{\mathrm{d}\ell_{\mathrm{p}}}{\mathrm{d}\boldsymbol{\beta}}\bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{\mathbb{B}}} + o_{\mathrm{p}}(1). \tag{20}$$

Inserting (20) into (17) and using (19) gives that

$$2\{\ell_{\mathrm{p}}(\hat{\boldsymbol{\beta}}) - \ell_{\mathrm{p}}(\hat{\boldsymbol{\beta}}_{\mathbb{B}})\} \xrightarrow{\mathrm{d}} \boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{\Sigma}\boldsymbol{H})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{I} - \boldsymbol{\Sigma}\boldsymbol{H})\boldsymbol{Z}, \quad n \to \infty.$$

It is straightforward to check that this has the same distribution as $\boldsymbol{W}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{\Sigma}^{1/2}\boldsymbol{H}\boldsymbol{\Sigma}^{1/2})\boldsymbol{W}$, where $\boldsymbol{W} \sim N(\boldsymbol{0}, \boldsymbol{I}_K)$ and $\boldsymbol{I} - \boldsymbol{\Sigma}^{1/2}\boldsymbol{H}\boldsymbol{\Sigma}^{1/2}$ is idempotent of rank m; thus its eigenvalues are 1 (m times) and 0 ($K - m$ times). Hence the limiting distribution of $2\{\ell_{\mathrm{p}}(\hat{\boldsymbol{\beta}}) - \ell_{\mathrm{p}}(\hat{\boldsymbol{\beta}}_{\mathbb{B}})\}$ is χ^2_m, as was announced. □

REFERENCES

Einmahl, J. H. J. and Segers, J. (2009), "Maximum Empirical Likelihood Estimation of the Spectral Measure of an Extreme-value Distribution," *The Annals of Statistics*, 37, 2953–2989.

Huang, A. and Rathouz, P. J. (2012), "Proportional Likelihood Ratio Models for Mean Regression," *Biometrika*, 99, 223–229.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979), *Multivariate Analysis*, London: Academic Press.

Owen, A. B. (2001), *Empirical Likelihood*, Boca Raton: Chapman & Hall.

Qin, J. and Lawless, J. (1994), "Empirical Likelihood and General Estimating Equations," *The Annals of Statistics*, 22, 300–325.