Supplementary materials for "Smoothing parameter selection in two frameworks for penalized splines"

Tatyana Krivobokova [*]
Georg-August-Universität Göttingen

12th October 2012

1 Estimating equations and their derivatives

In the following $\partial \boldsymbol{S}_\lambda/\partial\lambda = -\lambda^{-1}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)$ will be used.

1.1 Mallows' C_p

The Mallows' C_p is defined as

$$\mathrm{C}_p(\lambda) = \frac{1}{n}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{Y}\left\{1 + \frac{2\mathrm{tr}(\boldsymbol{S}_\lambda)}{n}\right\}.$$

Its estimation equation (obtained as $\lambda/2\ \partial \mathrm{C}_p(\lambda)/\partial\lambda$) will be denoted by $T_{Cp}(\lambda)$, which, together with its derivative, is given by

$$\begin{aligned}
T_{Cp}(\lambda) &= \frac{1}{n}\left[\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{Y}\left\{1 + \frac{2\mathrm{tr}(\boldsymbol{S}_\lambda)}{n}\right\} - \boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{Y}\ \frac{\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n}\right], \\
T'_{Cp}(\lambda) &= -\frac{1}{\lambda n}\left\{\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2(\boldsymbol{S}_\lambda - 3\boldsymbol{S}_\lambda^2)\boldsymbol{Y} - \frac{\mathrm{tr}(\boldsymbol{S}_\lambda)}{n}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2(\boldsymbol{I}_n - 6\boldsymbol{S}_\lambda + 6\boldsymbol{S}_\lambda^2)\boldsymbol{Y}\right. \\
&+ \left.\frac{\mathrm{tr}(\boldsymbol{S}_\lambda^2)}{n}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2(3\boldsymbol{I}_n - 4\boldsymbol{S}_\lambda)\boldsymbol{Y} - \frac{2\mathrm{tr}(\boldsymbol{S}_\lambda^3)}{n}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{Y}\right\}.
\end{aligned}$$

[*] Courant Research Center "Poverty, equity and growth" and Institute for Mathematical Stochastics, Georg-August-Universität Göttingen, Wilhelm-Weber-Str. 2, 37073 Göttingen, Germany

1.1.1 Frequentist model

Let find the expectations of $T_{Cp}(\lambda)$ and $T'_{Cp}(\lambda)$, as well as the variance of $T_{Cp}(\lambda)$, under the frequentist model (1), that is for $\boldsymbol{Y}$ with $\mathrm{E}_f(\boldsymbol{Y}) = \boldsymbol{f}$ and $\mathrm{var}_f(\boldsymbol{Y}) = \sigma^2 \boldsymbol{I}_n$.

$$
\begin{aligned}
\mathrm{E}_f\left\{T_{Cp}(\lambda)\right\} &= \frac{1}{n}\Big[\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f} - \sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2\} \\
&+ \frac{2\mathrm{tr}(\boldsymbol{S}_\lambda)}{n}\left\{\sigma^2\mathrm{tr}(2\boldsymbol{S}_\lambda - 3\boldsymbol{S}_\lambda^2 + \boldsymbol{S}_\lambda^3) + \boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f}\right\} \\
&- \frac{\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n}\left\{\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda^2) + \boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{f}\right\}\Big],
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}_f\left\{T'_{Cp}(\lambda)\right\} &= -\frac{1}{\lambda n}\Big[\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2(\boldsymbol{S}_\lambda - 3\boldsymbol{S}_\lambda^2)\boldsymbol{f} - \sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2(2\boldsymbol{I}_n - 3\boldsymbol{S}_\lambda)\} \\
&+ \frac{\mathrm{tr}(\boldsymbol{S}_\lambda)}{n}\left\{\sigma^2\mathrm{tr}(8\boldsymbol{S}_\lambda - 19\boldsymbol{S}_\lambda^2 + 18\boldsymbol{S}_\lambda^3 - 6\boldsymbol{S}_\lambda^4) - \boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2(\boldsymbol{I}_n - 6\boldsymbol{S}_\lambda + \boldsymbol{S}_\lambda^2)\boldsymbol{f}\right\} \\
&- \frac{\mathrm{tr}(\boldsymbol{S}_\lambda^2)}{n}\left\{\sigma^2\mathrm{tr}(10\boldsymbol{S}_\lambda - 11\boldsymbol{S}_\lambda^2 + 4\boldsymbol{S}_\lambda^3) - \boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2(3\boldsymbol{I}_n - 4\boldsymbol{S}_\lambda)\boldsymbol{f}\right\} \\
&+ \frac{2\mathrm{tr}(\boldsymbol{S}_\lambda^3)}{n}\left\{\sigma^2\mathrm{tr}(2\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2) - \boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{f}\right\}\Big].
\end{aligned}
$$

Under assumptions of Gaussian errors, one finds

$$
\begin{aligned}
\mathrm{var}_f\left\{T_{Cp}(\lambda)\right\} &= \frac{2\sigma^2}{n^2}\Big(\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda^2\} + 2\boldsymbol{f}(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda^2\boldsymbol{f} \\
&+ \frac{4\mathrm{tr}(\boldsymbol{S}_\lambda)}{n}\{1 + \mathrm{tr}(\boldsymbol{S}_\lambda)/n\}\left[\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda^2\} + 2\boldsymbol{f}(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda^2\boldsymbol{f}\right] \\
&- \frac{2\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n}\{1 + 2\mathrm{tr}(\boldsymbol{S}_\lambda)/n\}\left[\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda\} + 2\boldsymbol{f}(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda\boldsymbol{f}\right] \\
&+ \frac{\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)^2}{n^2}\{\sigma^2\mathrm{tr}(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4 + 2\boldsymbol{f}(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{f}\}\Big),
\end{aligned}
$$

using $\mathrm{var}_f(\boldsymbol{Y}^t\boldsymbol{A}\boldsymbol{Y}) = 2\sigma^4\mathrm{tr}(\boldsymbol{A}^2) + 4\sigma^2\boldsymbol{f}^t\boldsymbol{A}^2\boldsymbol{f}$ for any $n \times n$ matrix $\boldsymbol{A}$. If normality of errors is not given, but $\mathrm{E}_f(\epsilon_i^4) =: \mu_4 < \infty$ can be assumed, then $\mathrm{var}_f(\boldsymbol{Y}^t\boldsymbol{A}\boldsymbol{Y}) = 2\sigma^4\mathrm{tr}(\boldsymbol{A}^2) + 4\sigma^2\boldsymbol{f}^t\boldsymbol{A}^2\boldsymbol{f} + (\mu_4 - 3\sigma^4)\mathrm{tr}(\boldsymbol{A} \circ \boldsymbol{A})$, for $\circ$ denoting the Hadamard product (see e.g. Wiens, 1992). Hence, $\mathrm{var}_f\left\{T_{Cp}(\lambda)\right\}$ has an additional term, which, using the linearity of the Hadamard product, can be written as

$$
\begin{aligned}
& (\mu_4 - 3\sigma^4)\Big[\mathrm{tr}\left\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda \circ (\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\right\} \\
+ & \frac{4\mathrm{tr}(\boldsymbol{S}_\lambda)}{n}\{1 + 2\mathrm{tr}(\boldsymbol{S}_\lambda)/n\}\mathrm{tr}\left\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda \circ (\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\right\} \\
- & \frac{2\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n}\{1 + 2\mathrm{tr}(\boldsymbol{S}_\lambda)/n\}\mathrm{tr}\left\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda \circ (\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\right\} \\
+ & \frac{\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)^2}{n^2}\mathrm{tr}\left\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \circ (\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\right\}\Big].
\end{aligned}
$$

To simplify all above expressions, note that $\mathrm{tr}(\boldsymbol{S}_\lambda^l) = \mathrm{const}\lambda^{-1/(2q)}$ and $\lambda^{-1/(2q)}n^{-1} = o(1)$ due to (A3), as well as

$$
\begin{aligned}
\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\boldsymbol{f} &= \sum_{i=1}^{k+p+1}\frac{b_i^2(\lambda n\eta_i)^m}{(1+\lambda n\eta_i)^{m+l}} + n\ I_{\{l=0\}}\frac{1}{n}\boldsymbol{f}^t(\boldsymbol{I_n} - \boldsymbol{\Phi}_k\boldsymbol{\Phi}_k^t)\boldsymbol{f} \\
&= \lambda n\ \frac{1}{n}\sum_{i=1}^{k+p+1} b_i^2 n\eta_i\frac{(\lambda n\eta_i)^{m-1}}{(1+\lambda n\eta_i)^{m+l}} + O\left(k^{-2q}n\right) = O(\lambda n) + O\left(k^{-2q}n\right),
\end{aligned}
$$

for any $m = 1, 2, \dots$ and $l = 0, 1, \dots$. Here $n^{-1}\boldsymbol{f}^t(\boldsymbol{I_n} - \boldsymbol{\Phi}_k\boldsymbol{\Phi}_k^t)\boldsymbol{f}$ is the average squared approximation bias, see Claeskens et al. (2009).

Next, the diagonal elements $\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\}_{jj} = \sum_{i=q+1}^{k+p+1}\phi_{ji}^2(\lambda n\eta_i)^2(1+\lambda n\eta_i)^{-3}$, so that

$$
\begin{aligned}
\mathrm{tr}\left\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda \circ (\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\right\} &= \sum_{j=1}^{n}\left\{\sum_{i=q+1}^{k+p+1}\frac{\phi_{ji}^2(\lambda n\eta_i)^2}{(1+\lambda n\eta_i)^3}\right\}^2 \\
&\le \left[\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\}\right]^2\sum_{j=1}^{n}\{\max_i \phi_{ji}^2\}^2 = O\left(n^{-1}\lambda^{-1/q}\right),
\end{aligned}
$$

since $\phi_{ji}^2 = O(n^{-1})$ by definition. Similarly, one can show that the terms containing $\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\}_{jj} = 1 - \sum_{i=1}^{k+p+1}\phi_{ji}^2 + \sum_{i=q+1}^{k+p+1}\phi_{ji}^2(\lambda n\eta_i)^2(1+\lambda n\eta_i)^{-2}$ are also negligible. Hence, both for Gaussian and non-normal errors with $\mathrm{E}_f(\epsilon_i^4) < \infty$, it holds

$$
\begin{aligned}
\mathrm{E}_f\left\{T_{Cp}(\lambda)\right\} &= \frac{1}{n}\Big[\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f} - \sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2\} + o(1)\Big], \\
\mathrm{E}_f\left\{T_{Cp}'(\lambda)\right\} &= \frac{1}{\lambda n}\Big[\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2(2\boldsymbol{I}_n - 3\boldsymbol{S}_\lambda)\} - \boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2(\boldsymbol{S}_\lambda - 3\boldsymbol{S}_\lambda^2)\boldsymbol{f} + o(1)\Big], \\
\mathrm{var}_f\left\{T_{Cp}(\lambda)\right\} &= \frac{2\sigma^2}{n^2}\Big[2\boldsymbol{f}(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda^2\boldsymbol{f} + \sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda^2\} + o(1)\Big].
\end{aligned}
$$

Note that other popular criteria like generalized cross validation (GCV) by Craven and Wahba (1978) or Akaike information criterion (AIC Akaike, 1969) are asymptotically equivalent to Mallows' C_p, so that all subsequent results for Mallows' Cp hold also for these criteria. Indeed,

$$\begin{aligned}
\mathrm{GCV}(\lambda) &= n^{-1}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{Y}\{1 - \mathrm{tr}(\boldsymbol{S}_\lambda)/n\}^{-2} \\
&= n^{-1}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{Y}\{1 + 2\mathrm{tr}(\boldsymbol{S}_\lambda)/n + 3\mathrm{tr}(\boldsymbol{S}_\lambda)^2/n^2 + \ldots\} \\
\exp\{\mathrm{AIC}(\lambda)\} &= n^{-1}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{Y}\exp\{2\mathrm{tr}(\boldsymbol{S}_\lambda)/n\} \\
&= n^{-1}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{Y}\{1 + 2\mathrm{tr}(\boldsymbol{S}_\lambda)/n + 2\mathrm{tr}(\boldsymbol{S}_\lambda)^2/n^2 + \ldots\},
\end{aligned}$$

where $\mathrm{tr}(\boldsymbol{S}_\lambda)^2/n^2 = \mathrm{const}\ \lambda^{-1/q}n^{-2} = o\left(\mathrm{tr}(\boldsymbol{S}_\lambda)/n\right)$.

1.1.2 Stochastic model

To find expectations and variances under the stochastic model (4), that is for $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I}_n + \sigma_u^2\widetilde{\boldsymbol{\mathcal{R}}})$ ($\widetilde{\boldsymbol{\mathcal{R}}}$ is defined in the proof of Theorem 2), note that for any $m = 1, 2, \ldots$ and $l = 0, 1, \ldots$

$$\begin{aligned}
\mathrm{E}_\beta\left\{\boldsymbol{Y}^T(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\boldsymbol{Y}\right\} &= \sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\} + \sigma_u^2\mathrm{tr}\left\{\boldsymbol{C}\widetilde{\boldsymbol{D}}^-\boldsymbol{C}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\right\} \\
&\times \left[1 + \frac{\mathrm{tr}\left\{(\widetilde{\boldsymbol{\mathcal{R}}} - \boldsymbol{C}\widetilde{\boldsymbol{D}}^-\boldsymbol{C}^t)(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\right\}}{\mathrm{tr}\left\{\boldsymbol{C}\widetilde{\boldsymbol{D}}^-\boldsymbol{C}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\right\}}\right] \\
&= \sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\} \\
&+ \sigma_u^2\lambda n\left[\mathrm{tr}\left\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^{m-1}\boldsymbol{S}_\lambda^{l+1}\right\} - qI_{\{m=1\}}\right]\{1 + \widetilde{r}(l, m)\},
\end{aligned}$$

where

$$\mathrm{tr}\left\{\boldsymbol{C}\widetilde{\boldsymbol{D}}^-\boldsymbol{C}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\right\} = \lambda n\left[\mathrm{tr}\left\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^{m-1}\boldsymbol{S}_\lambda^{l+1}\right\} - qI_{\{m=1\}}\right]$$

has been used and

$$\widetilde{r}(l,m)=\frac{\mathrm{tr}\left\{(\widetilde{\boldsymbol{\mathcal{R}}}-\boldsymbol{C}\widetilde{\boldsymbol{D}}^{-}\boldsymbol{C}^t)(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\right\}}{\mathrm{tr}\left\{\boldsymbol{C}\widetilde{\boldsymbol{D}}^{-}\boldsymbol{C}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^m\boldsymbol{S}_\lambda^l\right\}}=\begin{cases}o(1), & l,m\in\mathbb{N}\\ o\left(\lambda^{-1+1/(2q)}\right), & l=0,\ m\in\mathbb{N}\end{cases},$$

is shown to hold below. With this,

$$\mathrm{E}_\beta\left\{T_{Cp}(\lambda)\right\} = \frac{1}{n}\Big[\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)o(1)+(\sigma_u^2\lambda n-\sigma^2)\mathrm{tr}(\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda^3)\Big]\{1+o(1)\},$$

$$\mathrm{E}_\beta\left\{T_{Cp}'(\lambda)\right\} = \frac{1}{\lambda n}\Big[\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda^3)+(\sigma_u^2\lambda n-\sigma^2)\mathrm{tr}\{(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)(3\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda)\}\Big]\{1+o(1)\},$$

$$\begin{aligned}\mathrm{var}_\beta\left\{T_{Cp}(\lambda)\right\} &= \frac{2\sigma^4}{n^2}\mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^2\}\{1+o(1)\}+\frac{2(\sigma_u^2\lambda n-\sigma^2)}{n^2}\\ &\times \Big[\sigma_u^2\lambda n\ \mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^4\}+\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^3(2\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\}\Big]\{1+o(1)\}.\end{aligned}$$

Since $\lambda_r=\sigma^2/(\sigma_u^2 n)=\lambda_{f|r}\{1+o(1)\}$, all terms with the multiplier $(\sigma_u^2\lambda_{f|r}n-\sigma^2)$ in $\mathrm{E}_\beta\left\{T_{Cp}(\lambda_{f|r})\right\}$, $\mathrm{E}_\beta\left\{T_{Cp}'(\lambda_{f|r})\right\}$ and $\mathrm{var}_\beta\left\{T_{Cp}(\lambda_{f|r})\right\}$ are asymptotically negligible.

Let now show the order of $\widetilde{r}(l,m)$. With the Demmler-Reinsch basis, one can represent $\widetilde{\boldsymbol{\mathcal{R}}}=\boldsymbol{\Phi}_n\mathrm{diag}(\boldsymbol{\eta}_n^-)\boldsymbol{\Phi}_n^t$, as well as $\boldsymbol{C}\widetilde{\boldsymbol{D}}^{-}\boldsymbol{C}^t=\boldsymbol{\Phi}_k\mathrm{diag}(\boldsymbol{\eta}_k^-)\boldsymbol{\Phi}_k^t$, where the eigenvalues $\boldsymbol{\eta}_k^-=(\boldsymbol{0}_q,\eta_{k,q+1}^{-1},\ldots,\eta_{k,k+p+1}^{-1})^t$, $\boldsymbol{\eta}_n^-=(\boldsymbol{0}_q,\eta_{n,q+1}^{-1},\ldots,\eta_{n,n}^{-1})^t$ and $\boldsymbol{\Phi}_n$ corresponds to the Demmler-Reinsch basis for $k=n$. Denote also $\boldsymbol{\Phi}_{kn}=\boldsymbol{\Phi}_k^t\boldsymbol{\Phi}_n$ a $(k+p+1)\times n$ semi-orthonormal matrix, such that $\boldsymbol{\Phi}_{kn}\boldsymbol{\Phi}_{kn}^t=\boldsymbol{I}_{k+p+1}$. With this,

$$\begin{aligned}\widetilde{r}(l,m) &= \frac{\mathrm{tr}\left[\left\{\boldsymbol{\Phi}_{kn}\mathrm{diag}(\boldsymbol{\eta}_n^-)\boldsymbol{\Phi}_{kn}^t-\mathrm{diag}(\boldsymbol{\eta}_k^-)\right\}\mathrm{diag}\left\{(\lambda n\boldsymbol{\eta}_k)^m(1+\lambda n\boldsymbol{\eta}_k)^{-(l+m)}\right\}\right]}{\lambda n\ \mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^{m-1}\boldsymbol{S}_\lambda^{l+1}\}}\\ &+ I_{\{l=0\}}\frac{\mathrm{tr}\{(\mathrm{diag}(\boldsymbol{\eta}_n^-)(\boldsymbol{I}_n-\boldsymbol{\Phi}_{kn}^t\boldsymbol{\Phi}_{kn})\}}{\lambda n\ \mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^{m-1}\boldsymbol{S}_\lambda^{l+1}\}}.\end{aligned}$$

First, consider the diagonal elements of $\{\boldsymbol{\Phi}_{kn}\mathrm{diag}(\boldsymbol{\eta}_n^-)\boldsymbol{\Phi}_{kn}^t-\mathrm{diag}(\boldsymbol{\eta}_k^-)\}$, which are given by $\left(\sum_{i=q+1}^n\{\boldsymbol{\Phi}_{kn}\}_{ij}^2\eta_{n,i}^{-1}-\eta_{k,j}^{-1}\right)$, $j=q+1,\ldots,k+p+1$. The elements $\{\boldsymbol{\Phi}_{kn}\}_{ij}$ belong to a product of two Demmler-Reinsch bases at the same set of x-values but based on a different number of knots: $\boldsymbol{\Phi}_k$ and $\boldsymbol{\Phi}_n$. From the properties of the Demmler-Reinsch basis it holds that $\eta_{k,i} = \eta_{n,i}\{1 + o(1)\}$, as well as $\phi_{k,i} = \phi_{n,i}\{1 + o(1)\}$, where $o(1)$ is independent of i, for $i = o\{n^{2/(2q+1)}\}$. Hence, $\{\boldsymbol{\Phi}_{kn}\}^2_{jj} = 1 + o(1)$ and $\{\boldsymbol{\Phi}_{kn}\}^2_{ij} = o(n^{-1})$, $i \neq j$. Let fix an index $j^* \propto n^{1/(2q)}$, so that $j^* = o\{n^{2/(2q+1)}\}$ is fulfilled. Then, for $j \leq j^*$

$$\sum_{i=q+1}^{n} \{\boldsymbol{\Phi}_{kn}\}^2_{ij}\eta^{-1}_{n,i} - \eta^{-1}_{k,j} = o(1) \sum_{i=q+1}^{n} \frac{(i-q)^{-2q}}{c(\rho)^{2q}} + o(1)\eta^{-1}_{k,j} = o(1)\eta^{-1}_{k,j},$$

implying

$$\begin{aligned} & \operatorname{tr}\left[\left\{\boldsymbol{\Phi}_{kn}\operatorname{diag}(\boldsymbol{\eta}^-_n)\boldsymbol{\Phi}^t_{kn} - \operatorname{diag}(\boldsymbol{\eta}^-_k)\right\}\operatorname{diag}\left\{(\lambda n\boldsymbol{\eta}_k)^m(1+\lambda n\boldsymbol{\eta}_k)^{-(l+m)}\right\}\right] \\ = \;& o(1)\lambda n \operatorname{tr}\{\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^{m-1}\boldsymbol{S}^{l+1}_\lambda\}, \end{aligned}$$

since for $j > j^*$ the sum components of the trace are negligible. Similarly,

$$\operatorname{tr}\{(\operatorname{diag}(\boldsymbol{\eta}^-_n)(\boldsymbol{I}_n - \boldsymbol{\Phi}^t_{kn}\boldsymbol{\Phi}_{kn})\} = \sum_{i=q+1}^{n}\eta^{-1}_{n,i}\left(1 - \sum_{j=q+1}^{k+p+1}\{\boldsymbol{\Phi}_{kn}\}^2_{ij}\right) = o\left(n^{-1}\right),$$

so that $\widetilde{r}(l,m) = o(1)$ for $l, m \in \mathbb{N}$ and $\widetilde{r}(0,m) = o(\lambda^{1-1/(2q)})$, $m \in \mathbb{N}$.

1.2 Restricted maximum likelihood

The likelihood function for the model (5) is given by

$$-2l(\boldsymbol{\beta}, \sigma^2, \sigma^2_u; \boldsymbol{Y}) = n\log\sigma^2 + \log|\boldsymbol{V}_\lambda| + \sigma^{-2}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})^t\boldsymbol{V}^{-1}_\lambda(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}),$$

with $\boldsymbol{V}_\lambda = \boldsymbol{I}_n + \sigma^2_u\boldsymbol{Z}\boldsymbol{D}^{-1}\boldsymbol{Z}^t/\sigma^2$. Plugging in $\widetilde{\boldsymbol{\beta}} = (\boldsymbol{X}^t\boldsymbol{V}^{-1}_\lambda\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{V}^{-1}_\lambda\boldsymbol{Y}$, leads to the profile likelihood for σ^2 and σ^2_u. However, σ^2 and $\lambda = \sigma^2/(n\sigma^2_u)$ are better to be estimated from the restricted profile likelihood (Patterson and Thompson, 1971)

$$\begin{aligned} -2l_r(\sigma^2, \lambda; \boldsymbol{Y}) &= -2l(\widetilde{\beta}, \sigma^2, \sigma^2_u; \boldsymbol{Y}) + \log|\sigma^2\boldsymbol{X}^t\boldsymbol{V}^{-1}_\lambda\boldsymbol{X}| \\ &= (n-q)\log\sigma^2 + \log|\boldsymbol{V}_\lambda||\boldsymbol{X}^t\boldsymbol{V}^{-1}_\lambda\boldsymbol{X}| + \frac{\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y}}{\sigma^2}. \end{aligned}$$

Since $\widehat{\sigma}^2_{ML} = (n-q)^{-1}\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y}$, the profile restricted likelihood for λ results in

$$-2l_p(\lambda;\boldsymbol{Y}) = (n-q)\log \boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y} + \log|\boldsymbol{V}_\lambda||\boldsymbol{X}^t\boldsymbol{V}_\lambda^{-1}\boldsymbol{X}|.$$

Its first derivative equals

$$\begin{aligned}\frac{\partial l_p(\lambda;\boldsymbol{Y})}{\partial\lambda} &= -\frac{n-q}{2\lambda}\,\frac{\boldsymbol{Y}^t(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\boldsymbol{Y}}{\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y}} - \frac{1}{2}\mathrm{tr}\left\{\left(\boldsymbol{X}\boldsymbol{V}_\lambda^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^t\frac{\partial\boldsymbol{V}_\lambda^{-1}}{\partial\lambda}\boldsymbol{X}\right\}\\ &- \frac{1}{2}\mathrm{tr}\left(\boldsymbol{V}_\lambda^{-1}\frac{\partial\boldsymbol{V}_\lambda}{\partial\lambda}\right) = -\frac{n-q}{2\lambda}\,\frac{\boldsymbol{Y}^t(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\boldsymbol{Y}}{\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y}} + \frac{\mathrm{tr}(\boldsymbol{S}_\lambda) - q}{2\lambda}.\end{aligned}$$

The estimating equation is now defined via

$$\begin{aligned}T_{ML}(\lambda) &= -\frac{2\lambda\,\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y}}{n(n-q)}\,\frac{\partial l_p(\lambda;\boldsymbol{Y})}{\partial\lambda}\\ &= \frac{1}{n}\left[\boldsymbol{Y}^t(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\boldsymbol{Y} - \boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y}\,\{\mathrm{tr}(\boldsymbol{S}_\lambda) - q\}/(n-q)\right].\end{aligned}$$

The first derivative of $T_{ML}(\lambda)$ is given by

$$\begin{aligned}T^{'}_{ML}(\lambda) &= -\frac{1}{\lambda n}\left\{\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda(\boldsymbol{I}_n - 2\boldsymbol{S}_\lambda)\boldsymbol{Y} + \boldsymbol{Y}^t(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\boldsymbol{Y}\frac{\mathrm{tr}(\boldsymbol{S}_\lambda) - q}{n-q}\right.\\ &- \left.\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{Y}\frac{\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n-q}\right\}.\end{aligned}$$

1.2.1 Frequentist model

Let now find the expectations of $T_{ML}(\lambda)$ and $T^{'}_{ML}(\lambda)$, as well as the variance of $T_{ML}(\lambda)$, under the frequentist model (1).

$$\begin{aligned}\mathrm{E}_f\{T_{ML}(\lambda)\} &= \frac{1}{n}\Big[\boldsymbol{f}^t(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\boldsymbol{f} - \sigma^2\left\{\mathrm{tr}(\boldsymbol{S}_\lambda^2) - q\right\}\\ &+ \frac{\mathrm{tr}(\boldsymbol{S}_\lambda) - q}{n-q}\left\{\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda) - \sigma^2 q - \boldsymbol{f}(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{f}\right\}\Big],\end{aligned}$$

$$
\begin{aligned}
\mathrm{E}_f\left\{T_{ML}^{'}(\lambda)\right\} &= \frac{1}{\lambda n}\Big[2\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda^3)-\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda(\boldsymbol{I}_n-2\boldsymbol{S}_\lambda)\boldsymbol{f} \\
&- \frac{\mathrm{tr}(\boldsymbol{S}_\lambda)-q}{n-q}\left\{\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)+\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda\boldsymbol{f}\right\} \\
&- \frac{\mathrm{tr}(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)}{n-q}\left\{\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda)-\sigma^2 q-\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{f}\right\}\Big].
\end{aligned}
$$

Under assumption of the Gaussian errors

$$
\begin{aligned}
\mathrm{var}_f\left\{T_{ML}(\lambda)\right\} &= \frac{2\sigma^2}{n^2}\Big(\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^2\}+2\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^2\boldsymbol{f} \\
&- \frac{2\{\mathrm{tr}(\boldsymbol{S}_\lambda)-q\}}{n-q}\left[\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\}+2\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f}\right] \\
&+ \frac{\{\mathrm{tr}(\boldsymbol{S}_\lambda)-q\}^2}{(n-q)^2}\left\{\sigma^2\mathrm{tr}(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2+2\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{f}\right\}\Big).
\end{aligned}
$$

If the errors are not Gaussian, but $\mathrm{E}_f(\epsilon_i^4)=\mu_4<\infty$ holds, then $\mathrm{var}_f\left\{T_{ML}(\lambda)\right\}$ has an extra term given by

$$
\begin{aligned}
&(\mu_4-3\sigma^4)\Big[\mathrm{tr}\left\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda\circ(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda\right\} \\
&- \frac{2\{\mathrm{tr}(\boldsymbol{S}_\lambda)-q\}}{n-q}\mathrm{tr}\left\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda\circ(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\right\} \\
&+ \frac{\{\mathrm{tr}(\boldsymbol{S}_\lambda)-q\}^2}{(n-q)^2}\mathrm{tr}\left\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\circ(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\right\}\Big].
\end{aligned}
$$

With the same arguments as in Section 1.1.1, one obtains

$$
\begin{aligned}
\mathrm{E}_f\left\{T_{ML}(\lambda)\right\} &= \frac{1}{n}\Big[\boldsymbol{f}^t(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)\boldsymbol{f}-\sigma^2\left\{\mathrm{tr}(\boldsymbol{S}_\lambda^2)-q\right\}+o(1)\Big], \\
\mathrm{E}_f\left\{T_{ML}^{'}(\lambda)\right\} &= \frac{1}{\lambda n}\Big[2\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda^3)-\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda(\boldsymbol{I}_n-2\boldsymbol{S}_\lambda)\boldsymbol{f}+o(1)\Big], \\
\mathrm{var}_f\left\{T_{ML}(\lambda)\right\} &= \frac{2\sigma^2}{n^2}\Big[2\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^2\boldsymbol{f}+\sigma^2\{\mathrm{tr}(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^2\}+o(1)\Big].
\end{aligned}
$$

Now, it is easy to verify equation (6) of the paper

$$
\begin{aligned}
R(\lambda) &= \mathrm{E}_f\left\{T_{ML}(\lambda)\right\}-\mathrm{E}_f\left\{T_{Cp}(\lambda)\right\}=\frac{1}{n}\Big[\boldsymbol{f}^t(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)\boldsymbol{f}-\sigma^2\left\{\mathrm{tr}(\boldsymbol{S}_\lambda^2)-q\right\}+o(1) \\
&- \boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f}-\sigma^2\mathrm{tr}\{(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2\}+o(1)\Big] \\
&= \frac{1}{n}\Big[\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2\boldsymbol{f}-\sigma^2\left\{\mathrm{tr}(\boldsymbol{S}_\lambda^3)-q\right\}+o(1)\Big].
\end{aligned}
$$

Note also $\mathrm{E}_f\left[\boldsymbol{Y}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2\boldsymbol{Y}-\sigma^2\{\mathrm{tr}(\boldsymbol{S}_\lambda^2)-q\}\right]=\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2\boldsymbol{f}-\sigma^2\{\mathrm{tr}(\boldsymbol{S}_\lambda^3)-q\}$.

1.2.2 Stochastic model

Applying results of Section 1.1.2 gives

$$\begin{aligned}
\mathrm{E}_\beta\left\{T_{ML}(\lambda)\right\} &= \frac{1}{n}\Big\{\mathrm{tr}(\boldsymbol{S}_\lambda)o(1)+(\sigma_u^2\lambda n-\sigma^2)\mathrm{tr}(\boldsymbol{S}_\lambda^2)\Big\}\{1+o(1)\},\\
\mathrm{E}_\beta\left\{T_{ML}^{'}(\lambda)\right\} &= \frac{\sigma^2}{\lambda n}\Big\{\mathrm{tr}(\boldsymbol{S}_\lambda^2)+(\sigma_u^2\lambda n-\sigma^2)\mathrm{tr}(2\boldsymbol{S}_\lambda^3-\boldsymbol{S}_\lambda^2)\Big\}\{1+o(1)\},\\
\mathrm{var}_\beta\left\{T_{ML}(\lambda)\right\} &= \frac{2}{n^2}\Big[\sigma^4\mathrm{tr}(\boldsymbol{S}_\lambda^2)-(\sigma_u^2\lambda n-\sigma^2)\sigma^2\mathrm{tr}\{\boldsymbol{S}_\lambda^3(2\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\}\\
&+ \sigma_u^2\lambda n(\sigma_u^2\lambda n-\sigma^2)\mathrm{tr}(\boldsymbol{S}_\lambda^4)\Big]\{1+o(1)\}.
\end{aligned}$$

2 Detailed proof of Theorem 3

2.1 Proof for $\widehat{\lambda}_f$

From the Taylor expansion $0=T_{Cp}(\widehat{\lambda}_f)=T_{Cp}(\lambda_f)+T_{Cp}^{'}(\widetilde{\lambda})(\widehat{\lambda}_f-\lambda_f)$, for some $\widetilde{\lambda}$ between $\widehat{\lambda}_f$ and λ_f, it holds $\widehat{\lambda}_f-\lambda_f=-T_{Cp}(\lambda_f)/T_{Cp}^{'}(\widetilde{\lambda})$.

Showing

$$\frac{T_{Cp}(\lambda_f)-\mathrm{E}_f\{T_{Cp}(\lambda_f)\}}{\sqrt{\mathrm{var}_f\{T_{Cp}(\lambda_f)\}}}\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\quad\text{and}\quad\frac{T_{Cp}^{'}(\widetilde{\lambda})}{\mathrm{E}_f\left\{T_{Cp}^{'}(\lambda_f)\right\}}\xrightarrow{\mathcal{P}}1,$$

would allow to apply Slutsky's lemma and to conclude that

$$\left(\frac{\widehat{\lambda}_f}{\lambda_f}-1\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\frac{\mathrm{var}_f\{T_{Cp}(\lambda_f)\}}{\left[\lambda_f\mathrm{E}_f\left\{T_{Cp}^{'}(\lambda_f)\right\}\right]^2}\right).$$

To find $\mathrm{E}_f\left\{T_{Cp}^{'}(\lambda_f)\right\}$ and $\mathrm{var}_f\{T_{Cp}(\lambda_f)\}$, Lemma 3 is applied to get

$$\begin{aligned}
\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2(\boldsymbol{S}_\lambda-3\boldsymbol{S}_\lambda^2)\boldsymbol{f}\big|_{\lambda=\lambda_f} &= -2\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda^3)\big|_{\lambda=\lambda_f}\{1+o(1)\},\\
\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^4\boldsymbol{S}_\lambda^2\boldsymbol{f}\big|_{\lambda=\lambda_f} &= o\left(\lambda_f^{-\frac{1}{2q}}\right),
\end{aligned}$$

so that

$$\begin{aligned} \mathrm{E}_f \left\{ T'_{Cp}(\lambda_f) \right\} &= \frac{\sigma^2}{\lambda_f n} \operatorname{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2(4\boldsymbol{I}_n - 3\boldsymbol{S}_\lambda)\}\big|_{\lambda=\lambda_f} \{1 + o(1)\}, \\ \mathrm{var}_f\{T_{Cp}(\lambda_f)\} &= \frac{2\sigma^4}{n^2} \operatorname{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4 \boldsymbol{S}_\lambda^2\}\big|_{\lambda=\lambda_f} \{1 + o(1)\}. \end{aligned}$$

Employing the formula

$$\operatorname{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^m \boldsymbol{S}_\lambda^l\} = \frac{\lambda^{-1/(2q)}}{c(\rho)} \frac{\Gamma\{m + 1/(2q)\}\Gamma\{l - 1/(2q)\}}{2q\Gamma(l + m)}\{1 + o(1)\},$$

as well as $\Gamma(1 + x) = x\Gamma(x)$ and $\Gamma\{1 - 1/(2q)\}\Gamma\{1/(2q)\}/(2q) = 1/\mathrm{sinc}\{\pi/(2q)\}$ allows to simplify

$$\begin{aligned} \mathrm{E}_f \left\{ T'_{Cp}(\lambda_f) \right\} &= \frac{\sigma^2 \lambda_f^{-1/(2q)-1}}{n\ c(\rho)} \frac{(2q-1)(4q+1)}{16q^3 \ \mathrm{sinc}\{\pi/(2q)\}}\{1 + o(1)\}, \\ \mathrm{var}_f\{T_{Cp}(\lambda_f)\} &= \frac{2\sigma^4 \lambda_f^{-1/(2q)}}{n^2\ c(\rho)} \frac{(2q-1)(2q+1)(4q+1)(6q+1)}{3840q^5 \ \mathrm{sinc}\{\pi/(2q)\}}\{1 + o(1)\}. \end{aligned}$$

Consider now

$$\begin{aligned} n\, T_{Cp}(\lambda_f) &= \left[\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \boldsymbol{S}_\lambda \boldsymbol{Y} \left\{1 + \frac{\operatorname{tr}(\boldsymbol{S}_\lambda)}{n}\right\} - \boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \boldsymbol{Y} \frac{\operatorname{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n}\right]\Bigg|_{\lambda=\lambda_f} \\ &= \left[\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \boldsymbol{S}_\lambda \boldsymbol{Y} - \widehat{\sigma}^2 \operatorname{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\right]\big|_{\lambda=\lambda_f} + o_p(1). \end{aligned}$$

One can also represent,

$$\mathrm{E}_f\{n\ T_{Cp}(\lambda_f)\} = \left[\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \boldsymbol{S}_\lambda \boldsymbol{f} + \sigma^2 \operatorname{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \boldsymbol{S}_\lambda\} - \sigma^2 \operatorname{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\right]\big|_{\lambda=\lambda_f} + o(1).$$

Denoting $d_i = \sum_{j=1}^n \phi_{k,i}(x_j) y_j$, such that $\mathrm{E}_f(d_i^2) = b_i^2 + \sigma^2$, and noting that $\widehat{\sigma}^2 = \sigma^2\{1 + O_p(n^{-1/2})\}$, define random variables ξ_i

$$n\left[T_{Cp}(\lambda_f) - \mathrm{E}_f\{T_{Cp}(\lambda_f)\}\right] = \sum_{i=q+1}^{k+p+1} \left(d_i^2 - b_i^2 - \sigma^2\right) \frac{(\lambda_f n \eta_i)^2}{(1 + \lambda_f n \eta_i)^3} + o_p(1) =: \sum_{i=q+1}^{k+p+1} \xi_i,$$

such that $\mathrm{E}_f(\xi_i) = o(1)$ and $s_n^2 = \sum_{i=q+1}^{k+p+1} \mathrm{var}_f(\xi_i) = 2\sigma^4 \operatorname{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^4 \boldsymbol{S}_\lambda^2\}\{1 + o(1)\}$. Since $s_n^2 = \mathrm{const}\ \lambda_f^{-1/(2q)}$ and $(\lambda_f^{1/(2q)} k)^{-1} \to 0$ according to (A2) and (A3), each $\mathrm{var}_f(\xi_i) = o(1)$

and there exist a constant B, such that $\mathrm{E}_f|\xi_i|^2 = \mathrm{var}_f(\xi_i)+o(1) < B$, $i = q+1,\ldots,k+p+1$. With this, the Lyapunov's condition

$$s_n^{-4}\sum_{i=q+1}^{k+p+1}\mathrm{E}_f|\xi_i|^4 < Bs_n^{-4}\sum_{i=q+1}^{k+p+1}\mathrm{E}_f|\xi_i|^2 = Bs_n^{-2} = O\left(\lambda_f^{1/(2q)}\right)$$

converges to zero as n tends to infinity. Thus, $s_n^{-1}\sum_{i=q+1}^{k+p+1}\xi_i \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$, or equivalently, $[\mathrm{var}_f\{T_{Cp}(\lambda_f)\}]^{-1/2}[T_{Cp}(\lambda_f) - \mathrm{E}_f\{T_{Cp}(\lambda_f)\}] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$.

Next is shown that $\widehat{\lambda}_f \xrightarrow{\mathcal{P}} \lambda_f$. From $\mathrm{var}_f\{T_{Cp}(\lambda)\} = O\left(\lambda^{-1/(2q)}n^{-2}\right) \to 0$ for $n \to \infty$, it follows $T_{Cp}(\lambda) \xrightarrow{\mathcal{P}} \mathrm{E}_f\{T_{Cp}(\lambda)\}$, for any λ satisfying (A3). It remains to verify that $\mathrm{E}_f[T_{Cp}\{\lambda_f(1-\varepsilon)\}] < 0 < \mathrm{E}_f[T_{Cp}\{\lambda_f(1+\varepsilon)\}]$, for any $\varepsilon \in (0,1)$ (see Lemma 5.10 in van der Vaart, 1998). Let define $B_1(\lambda_f)$ and $B_2(\lambda_f)$ from the representation of $\mathrm{E}_f\{T_{Cp}(\lambda_f)\}$ in terms of the Demmler-Reinsch basis.

$$\begin{aligned}
\mathrm{E}_f\{T_{Cp}(\lambda_f)\} &= \frac{1}{n}\left[\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f} - \sigma^2\mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda^2\} + o(1)\right]\Big|_{\lambda=\lambda_f}\\
&= \frac{1}{n}\sum_{i=1}^{k+p+1}\frac{b_i^2(\lambda_f n\eta_i)^2}{(1+\lambda_f n\eta_i)^3} - \frac{\sigma^2\lambda_f^{-\frac{1}{2q}}c(q,2,K_q)}{4qc(\rho)n} + o\left(n^{-1}\right)\\
&=: B_1(\lambda_f) - B_2(\lambda_f) + o\left(n^{-1}\right),
\end{aligned}$$

where $B_1(\lambda_f) - B_2(\lambda_f) = 0$ by definition of λ_f. Then,

$$\begin{aligned}
\mathrm{E}_f[T_{Cp}\{\lambda_f(1-\varepsilon)\}] &= \frac{1}{n}\sum_{i=1}^{k+p+1}\frac{b_i^2\{\lambda_f(1-\varepsilon)n\eta_i\}^2}{\{1+\lambda_f n(1-\varepsilon)\eta_i\}^3}\\
&- (1-\varepsilon)^{-\frac{1}{2q}}\frac{\sigma^2\lambda_f^{-1/(2q)}c(q,2,K_q)}{4qnc(\rho)} + o\left(n^{-1}\right)\\
&= \frac{(1-\varepsilon)^2}{n}\left\{\sum_{i=1}^{k+p+1}\frac{b_i^2(\lambda_f n\eta_i)^2}{(1+\lambda_f n\eta_i)^3} + \sum_{j=1}^{\infty}\sum_{i=1}^{k+p+1}\frac{(j+2)(j+1)b_i^2(\lambda_f n\eta_i)^{2+j}}{\varepsilon^{-j}2(1+\lambda_f n\eta_i)^{3+j}}\right\}\\
&- (1-\varepsilon)^{-\frac{1}{2q}}B_2(\lambda_f) + o\left(n^{-1}\right)\\
&= (1-\varepsilon)^2 B_1(\lambda_f)\left\{1 + \sum_{j=1}^{\infty}\frac{\varepsilon^j(j+2)(j+1)}{2}\frac{\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^{2+j}\boldsymbol{S}_\lambda\boldsymbol{f}}{\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f}}\Bigg|_{\lambda=\lambda_f}\right\}\\
&- (1-\varepsilon)^{-\frac{1}{2q}}B_2(\lambda_f) + o\left(n^{-1}\right).
\end{aligned}$$

Since $\sum_{j=1}^{\infty}\varepsilon^j(j+1)(j+2)=2\varepsilon(\varepsilon^2-3\varepsilon+3)(1-\varepsilon)^{-3}$ and according to Lemma 3 it holds that $\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^{2+1}\boldsymbol{S}_\lambda\boldsymbol{f}\big|_{\lambda=\lambda_f}=o(1)\ \boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{f}\big|_{\lambda=\lambda_f}$, one gets

$$\begin{aligned}\mathrm{E}_f[T_{Cp}\{\lambda_f(1-\varepsilon)\}] &= (1-\varepsilon)^2 B_1(\lambda_f)\{1+o(1)\}-(1-\varepsilon)^{-\frac{1}{2q}}\,B_2(\lambda_f)\\ &= (1-\varepsilon)^2 B_1(\lambda_f)\left\{1-(1-\varepsilon)^{-2-\frac{1}{2q}}+o(1)\right\}<0,\end{aligned}$$

for $n\to\infty$. Similarly,

$$\mathrm{E}_f\left[T_{Cp}\{\lambda_f(1+\varepsilon)\}\right]=(1+\varepsilon)^2 B_1(\lambda_f)\left\{1-(1+\varepsilon)^{-2-\frac{1}{2q}}+o(1)\right\}>0,$$

for $n\to\infty$, so that $\widehat{\lambda}_f\xrightarrow{\mathcal{P}}\lambda_f$ follows.

Let now consider $T'_{Cp}(\tau\lambda_f)/\mathrm{E}_f\{T'_{Cp}(\lambda_f)\}$, where $\tau\in[1-\varepsilon,1+\varepsilon]$ for any bounded $\varepsilon>0$. It is easy to see that, since $\mathrm{var}_f\left\{T'_{Cp}(\tau\lambda_f)\right\}=(\tau\lambda_f)^{-2-1/(2q)}n^{-2}\mathrm{const}\{1+o(1)\}$,

$$\mathrm{var}_f\left[\frac{T'_{Cp}(\tau\lambda_f)}{\mathrm{E}_f\left\{T'_{Cp}(\lambda_f)\right\}}\right]=O\left(\lambda_f^{1/(2q)}\right)\to 0,\quad n\to\infty.$$

Also, using Lemma 3 and the same arguments as in the proof of $\widehat{\lambda}_f\xrightarrow{\mathcal{P}}\lambda_f$,

$$\begin{aligned}\mathrm{E}_f\left\{T'_{Cp}(\tau\lambda_f)\right\} &= \frac{\sigma^2}{\lambda_f\tau n}\left[\tau^{-\frac{1}{2q}}\mathrm{tr}\{(\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda^3)(2\boldsymbol{I}_n-3\boldsymbol{S}_\lambda)\}+\tau^2 2\mathrm{tr}(\boldsymbol{S}_\lambda^2-\boldsymbol{S}_\lambda^3)\right]\Big|_{\lambda=\lambda_f}\{1+o(1)\}\\ &= \mathrm{E}_f\{T'_{Cp}(\lambda_f)\}\ \frac{4q\,\tau+\tau^{-1-1/(2q)}}{4q+1}\{1+o(1)\},\end{aligned}$$

where

$$\frac{4q\,\tau+\tau^{-1-1/(2q)}}{4q+1}=\begin{cases}1-\varepsilon\{1-1/(2q)\}+O(\varepsilon^2), & \text{for }\ \tau=1-\varepsilon\\ 1+\varepsilon\{1-1/(2q)\}+O(\varepsilon^2), & \text{for }\ \tau=1+\varepsilon\end{cases},$$

so that for any fixed $\tau\in[1-\varepsilon,1+\varepsilon]$ it holds $T'_{Cp}(\tau\lambda_f)/\mathrm{E}_f\{T'_{Cp}(\lambda_f)\}\xrightarrow{\mathcal{P}}1$, as $n\to\infty$. Since $P(|\widetilde{\lambda}/\lambda_f-1|\le\varepsilon)\to 1$ for $n\to\infty$ and any $\varepsilon>0$ due to $\widehat{\lambda}_f\xrightarrow{\mathcal{P}}\lambda_f$, it follows

$$\frac{T'_{Cp}(\widetilde{\lambda})}{\mathrm{E}_f\left\{T'_{Cp}(\lambda_f)\right\}}\xrightarrow{\mathcal{P}}1,\quad n\to\infty.$$

Putting all together and applying Slutsky's lemma gives

$$\left(\frac{\widehat{\lambda}_f}{\lambda_f}-1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\lambda_f^{1/(2q)} c(\rho)\mathrm{sinc}\{\pi/(2q)\}\frac{q(12q^2+8q+1)}{15(8q^2-2q-1)}\right).$$

2.2 Proof for $\widehat{\lambda}_{r|f}$

Proof for $\widehat{\lambda}_{r|f}$ follows the same lines, using equations derived in Section 1.2. Consider the first order Taylor expansion $0 = T_{ML}(\widehat{\lambda}_r) = T_{ML}(\lambda_{r|f}) + T'_{ML}(\widetilde{\lambda})(\widehat{\lambda}_r - \lambda_{r|f})$, for some $\widetilde{\lambda}$ between $\widehat{\lambda}_r$ and $\lambda_{r|f}$ and show that

$$\frac{T_{ML}(\lambda_{r|f}) - \mathrm{E}_f\{T_{ML}(\lambda_{r|f})\}}{\sqrt{\mathrm{var}_f\{T_{ML}(\lambda_{r|f})\}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \text{and} \quad \frac{T'_{ML}(\widetilde{\lambda})}{\mathrm{E}_f\left\{T'_{ML}(\lambda_{r|f})\right\}} \xrightarrow{\mathcal{P}} 1.$$

Applying Lemma 3 to see that

$$\begin{aligned}
\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda(\boldsymbol{I}_n-2\boldsymbol{S}_\lambda)\boldsymbol{f}\big|_{\lambda=\lambda_{r|f}} &= -\sigma^2\left\{\mathrm{tr}(\boldsymbol{S}_\lambda^2)-q\right\}\big|_{\lambda=\lambda_{r|f}}\{1+o(1)\},\\
\boldsymbol{f}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda^2\boldsymbol{f}\big|_{\lambda=\lambda_{r|f}} &= o\left(\lambda_{r|f}^{-1/(2q)}\right),
\end{aligned}$$

and simplifying Gamma functions results in

$$\begin{aligned}
\mathrm{E}_f\left\{T'_{ML}(\lambda_{r|f})\right\} &= \frac{\sigma^2\lambda_{r|f}^{-1/(2q)-1}}{n\ c(\rho)}\frac{4q^2-1}{4q^2\ \mathrm{sinc}\{\pi/(2q)\}}\{1+o(1)\},\\
\mathrm{var}_f\left\{T_{ML}(\lambda_{r|f})\right\} &= \frac{2\sigma^4\lambda_{r|f}^{-1/(2q)}}{n^2\ c(\rho)}\frac{4q^2-1}{48q^3\ \mathrm{sinc}\{\pi/(2q)\}}\{1+o(1)\}.
\end{aligned}$$

Consider now

$$\begin{aligned}
n\,T_{ML}(\lambda_{r|f}) &= \left[\boldsymbol{Y}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda\boldsymbol{Y} - \frac{\boldsymbol{Y}^t(\boldsymbol{I}_n-\boldsymbol{S}_\lambda)\boldsymbol{Y}}{n-q}\{\mathrm{tr}(\boldsymbol{S}_\lambda)-q\}\right]\bigg|_{\lambda=\lambda_{r|f}}\\
&= \left[\boldsymbol{Y}^t(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)\boldsymbol{Y}-\sigma^2\{\mathrm{tr}(\boldsymbol{S}_\lambda)-q\}\right]\big|_{\lambda=\lambda_{r|f}}+o_p(1)\\
\mathrm{E}_f\left\{n\,T_{ML}(\lambda_{r|f})\right\} &= \left[\boldsymbol{f}^t(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)\boldsymbol{f}+\sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda-\boldsymbol{S}_\lambda^2)-\sigma^2\{\mathrm{tr}(\boldsymbol{S}_\lambda)-q\}\right]\big|_{\lambda=\lambda_{r|f}}+o(1).
\end{aligned}$$

Define random variables ξ_i by

$$n\left[T_{ML}(\lambda_{r|f}) - \mathrm{E}_f\{T_{ML}(\lambda_{r|f})\}\right] = \sum_{i=q+1}^{k+p+1} \left(d_i^2 - b_i^2 - \sigma^2\right) \frac{\lambda_f n \eta_i}{(1+\lambda_f n\eta_i)^2} + o_p(1) =: \sum_{i=q+1}^{k+p+1} \xi_i,$$

such that $\mathrm{E}_f(\xi_i) = o(1)$ and $s_n^2 = \sum_{i=q+1}^{k+p+1} \mathrm{var}_f(\xi_i) = 2\sigma^4 \mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \boldsymbol{S}_\lambda^2\}\{1+o(1)\}$. Since $s_n^2 = \text{const}\ \lambda_f^{-1/(2q)}$ and $(\lambda_f^{1/(2q)} k)^{-1} \to 0$, according to (A2) and (A3), each $\mathrm{var}_f(\xi_i) = o(1)$ and there exist a constant B, such that $\mathrm{E}_f|\xi_i|^2 = \mathrm{var}_f(\xi_i) + o(1) < B$, $i = q+1, \ldots, k+p+1$. With this, the Lyapunov's condition

$$s_n^{-4} \sum_{i=q+1}^{k+p+1} \mathrm{E}_f|\xi_i|^4 < B s_n^{-4} \sum_{i=q+1}^{k+p+1} \mathrm{E}_f|\xi_i|^2 = B s_n^{-2} = O\left(\lambda_{r|f}^{1/(2q)}\right)$$

converges to 0, $n \to \infty$ and $\left[\mathrm{var}_f\{T_{ML}(\lambda_{r|f})\}\right]^{-1/2} \left[T_{ML}(\lambda_{r|f}) - \mathrm{E}_f\{T_{ML}(\lambda_{r|f})\}\right] \stackrel{\mathcal{D}}{\to} \mathcal{N}(0,1)$. Next is shown that $\widehat{\lambda}_r \stackrel{\mathcal{P}}{\to} \lambda_{r|f}$. From $\mathrm{var}_f\{T_{ML}(\lambda)\} = O\left(\lambda^{-1/(2q)} n^{-2}\right) \to 0$ for $n \to \infty$, it follows $T_{ML}(\lambda) \stackrel{\mathcal{P}}{\to} \mathrm{E}_f\{T_{ML}(\lambda)\}$, for any λ satisfying (A3). It remains to verify that $\mathrm{E}_f\left[T_{ML}\{\lambda_{r|f}(1-\varepsilon)\}\right] < 0 < \mathrm{E}_f\left[T_{ML}\{\lambda_{r|f}(1+\varepsilon)\}\right]$ for $\varepsilon \in (0,1)$. Define $B_1(\lambda_{r|f})$ and $B_2(\lambda_{r|f})$ from the representation of $\mathrm{E}_f\{T_{ML}(\lambda_{r|f})\}$ in terms of the Demmler-Reinsch basis.

$$\begin{aligned}
\mathrm{E}_f\{T_{ML}(\lambda_{r|f})\} &= \frac{1}{n}\left[\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda \boldsymbol{f} - \sigma^2\{\mathrm{tr}(\boldsymbol{S}_\lambda^2) - q\} + o(1)\right]\Big|_{\lambda=\lambda_{r|f}} \\
&= \frac{1}{n}\sum_{i=1}^{k+p+1} \frac{b_i^2 \lambda_{r|f} n\eta_i}{(1+\lambda_{r|f} n\eta_i)^2} - \frac{\sigma^2 \lambda_{r|f}^{-1/(2q)} c(q,2,K_q)}{c(\rho) n} + o\left(n^{-1}\right) \\
&=: B_1(\lambda_{r|f}) - B_2(\lambda_{r|f}) + o\left(n^{-1}\right),
\end{aligned}$$

where $B_1(\lambda_{r|f}) - B_2(\lambda_{r|f}) = o\left(n^{-1}\right)$ by definition of $\lambda_{r|f}$. Then,

$$\begin{aligned}
\mathrm{E}_f[T_{ML}\{\lambda_{r|f}(1-\varepsilon)\}] &= \frac{1}{n}\sum_{i=1}^{k+p+1} \frac{b_i^2 \lambda_{r|f}(1-\varepsilon) n\eta_i}{\{1+\lambda_{r|f} n(1-\varepsilon)\eta_i\}^2} \\
&- (1-\varepsilon)^{-\frac{1}{2q}} \frac{2\sigma^2 \lambda_{r|f}^{-1/(2q)} c(q,2,K_q)}{nc(\rho)} + o\left(n^{-1}\right) \\
&= (1-\varepsilon)^2 B_1(\lambda_{r|f}) \left\{1 + \sum_{j=1}^{\infty} (j+1)\varepsilon^j \left. \frac{\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^{1+j} \boldsymbol{S}_\lambda \boldsymbol{f}}{\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda \boldsymbol{f}} \right|_{\lambda=\lambda_{r|f}} \right\} \\
&- (1-\varepsilon)^{-\frac{1}{2q}} B_2(\lambda_{r|f}) + o\left(n^{-1}\right).
\end{aligned}$$

Since $\sum_{j=1}^{\infty}(j+1)\varepsilon^j = \varepsilon(2-\varepsilon)(1-\varepsilon)^{-2}$, and according to Lemma 3 it holds that $\boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^{1+1}\boldsymbol{S}_\lambda \boldsymbol{f}\big|_{\lambda=\lambda_{r|f}} = o(1)\ \boldsymbol{f}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)\boldsymbol{S}_\lambda \boldsymbol{f}\big|_{\lambda=\lambda_{r|f}}$, one gets

$$\begin{aligned} \mathrm{E}_f[T_{ML}\{\lambda_{r|f}(1-\varepsilon)\}] &= (1-\varepsilon)^2 B_1(\lambda_{r|f})\{1+o(1)\} - (1-\varepsilon)^{-\frac{2q+1}{2q}}\ B_2(\lambda_{r|f}) \\ &= (1-\varepsilon)^2 B_1(\lambda_{r|f})\left\{1-(1-\varepsilon)^{-2-\frac{1}{2q}} + o(1)\right\} < 0, \end{aligned}$$

for $n \to \infty$. Similarly,

$$\mathrm{E}_f\left[T_{ML}\{\lambda_{r|f}(1+\varepsilon)\}\right] = (1+\varepsilon)^2 B_1(\lambda_{r|f})\left\{1-(1+\varepsilon)^{-2-\frac{1}{2q}} + o(1)\right\} > 0,$$

for $n \to \infty$, so that $\widehat{\lambda}_r \xrightarrow{\mathcal{P}} \lambda_{r|f}$ follows.

Let now consider $T'_{ML}(\tau\lambda_{r|f})/\mathrm{E}_f\{T'_{ML}(\lambda_{r|f})\}$, where $\tau \in [1-\varepsilon, 1+\varepsilon]$, for any bounded $\varepsilon > 0$. It is easy to see that, since $\mathrm{var}_f\left\{T'_{ML}(\tau\lambda_{r|f})\right\} = (\tau\lambda_{r|f})^{-2-1/(2q)} n^{-2}\mathrm{const}\{1+o(1)\}$,

$$\mathrm{var}_f\left[\frac{T'_{ML}(\tau\lambda_{r|f})}{\mathrm{E}_f\left\{T'_{ML}(\lambda_{r|f})\right\}}\right] = O\left(\lambda_{r|f}^{1/(2q)}\right) \to 0, \quad n \to \infty.$$

Also, using Lemma 3 and the same arguments as in the proof of $\widehat{\lambda}_r \xrightarrow{\mathcal{P}} \lambda_{r|f}$,

$$\begin{aligned} \mathrm{E}_f\left\{T'_{ML}(\tau\lambda_{r|f})\right\} &= \frac{\sigma^2}{\lambda_{r|f}\tau n}\left[\tau^{-\frac{1}{2q}} 2\mathrm{tr}(\boldsymbol{S}_\lambda^2 - \boldsymbol{S}_\lambda^3) + \tau^2\{\mathrm{tr}(\boldsymbol{S}_\lambda^2) - q\}\right]\Big|_{\lambda=\lambda_{r|f}}\{1+o(1)\} \\ &= \mathrm{E}_f\{T'_{ML}(\lambda_{r|f})\}\ \frac{2q\,\tau + \tau^{-1-1/(2q)}}{2q+1}\{1+o(1)\}, \end{aligned}$$

where

$$\frac{2q\,\tau + \tau^{-1-1/(2q)}}{2q+1} = \begin{cases} 1-\varepsilon\{1-1/(2q)-1/(2q+1)\} + O(\varepsilon^2), & \text{for } \ \tau = 1-\varepsilon \\ 1+\varepsilon\{1-1/(2q)-1/(2q+1)\} + O(\varepsilon^2), & \text{for } \ \tau = 1+\varepsilon \end{cases},$$

so that for any fixed $\tau \in [1-\varepsilon, 1+\varepsilon]$ it holds $T'_{ML}(\tau\lambda_{r|f})/\mathrm{E}_f\{T'_{ML}(\lambda_{r|f})\} \xrightarrow{\mathcal{P}} 1$, as $n \to \infty$. Since $P(|\widetilde{\lambda}/\lambda_{r|f} - 1| \leq \varepsilon) \to 1$ for $n \to \infty$ and any $\varepsilon > 0$ due to $\widehat{\lambda}_r \xrightarrow{\mathcal{P}} \lambda_{r|f}$, it follows

$$\frac{T'_{ML}(\widetilde{\lambda})}{\mathrm{E}_f\left\{T'_{ML}(\lambda_{r|f})\right\}} \xrightarrow{\mathcal{P}} 1.$$

Putting all together and applying Slutsky's lemma gives

$$\left(\frac{\widehat{\lambda}_r}{\lambda_{r|f}} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\lambda_{r|f}^{1/(2q)} c(\rho)\mathrm{sinc}\{\pi/(2q)\}\frac{q}{12q^2-3}\right).$$

3 Detailed proof of Theorem 4

3.1 Proof for $\widehat{\lambda}_f$

All the steps of the proof are the same as in Theorem 3, that is one needs to show

$$\frac{T_{Cp}(\lambda_{f|r}) - \mathrm{E}_\beta\{T_{Cp}(\lambda_{f|r})\}}{\sqrt{\mathrm{var}_\beta\{T_{Cp}(\lambda_{f|r})\}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \text{and} \quad \frac{T'_{Cp}(\widetilde{\lambda})}{\mathrm{E}_\beta\left\{T'_{Cp}(\lambda_{f|r})\right\}} \xrightarrow{\mathcal{P}} 1,$$

for some $\widetilde{\lambda}$ between $\widehat{\lambda}_f$ and $\lambda_{f|r}$. Simplifying Gamma functions in the expressions for $\mathrm{E}_\beta\left\{T'_{Cp}(\lambda_{f|r})\right\}$ and $\mathrm{var}_\beta\left\{T_{Cp}(\lambda_{f|r})\right\}$ obtained in Section 1.1.2, results in

$$\begin{aligned}
\mathrm{E}_\beta\left\{T'_{Cp}(\lambda_{f|r})\right\} &= \frac{\sigma^2\lambda_{f|r}^{-1/(2q)-1}}{n\ c(\rho)}\frac{2q-1}{8q^2\ \mathrm{sinc}\{\pi/(2q)\}}\{1+o(1)\},\\
\mathrm{var}_\beta\left\{T_{Cp}(\lambda_{f|r})\right\} &= \frac{2\sigma^4\lambda_{f|r}^{-1/(2q)}}{n^2\ c(\rho)}\frac{4q^2-1}{48q^3\ \mathrm{sinc}\{\pi/(2q)\}}\{1+o(1)\}.
\end{aligned}$$

Consider now

$$\begin{aligned}
n\ T_{Cp}(\lambda_{f|r}) &= \left[\boldsymbol{Y}^t(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2\boldsymbol{S}_\lambda\boldsymbol{Y} - \sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\right]\big|_{\lambda=\lambda_{f|r}} + o_p(1),\\
\mathrm{E}_\beta\left\{n\ T_{Cp}(\lambda_{f|r})\right\} &= \sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\big|_{\lambda=\lambda_{f|r}}\ o(1).
\end{aligned}$$

For $d_i = \sum_{j=1}^n \phi_{k,i}(x_j)y_j$, such that $\mathrm{E}_\beta(d_i^2) = \sigma^2\lambda_{f|r}n\eta_i(1+\lambda_{f|r}n\eta_i)^{-1}\{1+o(1)\}$, let define random variables ξ_i by

$$\begin{aligned}
n\ \left[T_{Cp}(\lambda_{f|r}) - \mathrm{E}_f\{T_{Cp}(\lambda_{f|r})\}\right] &= \sum_{i=q+1}^{k+p+1}\left[d_i^2 - \sigma^2\frac{1+\lambda_{f|r}n\eta_i}{\lambda_{f|r}n\eta_i}\{1+o(1)\}\right]\frac{(\lambda_{f|r}n\eta_i)^2}{(1+\lambda_{f|r}n\eta_i)^3} + o_p(1)\\
&=: \sum_{i=q+1}^{k+p+1}\xi_i,
\end{aligned}$$

with $\mathrm{E}_\beta(\xi_i) = o(1)$ and $s_n^2 = \sum_{i=q+1}^{k+p+1} \mathrm{var}_\beta(\xi_i) = 2\sigma^4 \mathrm{tr}\{(\boldsymbol{I}_n - \boldsymbol{S}_\lambda)^2 \boldsymbol{S}_\lambda^2\}\{1+o(1)\}$. Since $s_n^2 = \text{const}\ \lambda_{f|r}^{-1/(2q)}$ and $(\lambda_{f|r}^{1/(2q)} k)^{-1} \to 0$ according to (A2), each $\mathrm{var}_\beta(\xi_i) = o(1)$ and there exist a constant B, such that $\mathrm{E}_\beta|\xi_i|^2 = \mathrm{var}_\beta(\xi_i) + o(1) < B$, $i = q+1, \ldots, k+p+1$. With this, the Lyapunov's condition

$$s_n^{-4} \sum_{i=q+1}^{k+p+1} \mathrm{E}_\beta|\xi_i|^4 < B s_n^{-4} \sum_{i=q+1}^{k+p+1} \mathrm{E}_\beta|\xi_i|^2 = B s_n^{-2} = O\left(\lambda_{f|r}^{1/(2q)}\right)$$

converges to zero as n tends to infinity. Thus, $s_n^{-1} \sum_{i=q+1}^{k+p+1} \xi_i \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$, or equivalently, $\left[\mathrm{var}_\beta\{T_{Cp}(\lambda_{f|r})\}\right]^{-1/2} \left[T_{Cp}(\lambda_{f|r}) - \mathrm{E}_\beta\{T_{Cp}(\lambda_{f|r})\}\right] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$.

Next is shown that $\widehat{\lambda}_f \xrightarrow{\mathcal{P}} \lambda_{f|r}$. From $\mathrm{var}_\beta\{T_{Cp}(\lambda)\} = O\left(\lambda^{-1/(2q)} n^{-2}\right) \to 0$, for $n \to \infty$ it follows $T_{Cp}(\lambda) \xrightarrow{\mathcal{P}} \mathrm{E}_\beta\{T_{Cp}(\lambda)\}$. It remains to verify that $\mathrm{E}_\beta\left[T_{Cp}\{\lambda_{f|r}(1-\varepsilon)\}\right] < 0 < \mathrm{E}_\beta\left[T_{Cp}\{\lambda_{f|r}(1+\varepsilon)\}\right]$, for any $\varepsilon \in (0,1)$. Indeed,

$$\begin{aligned}
\mathrm{E}_\beta[T_{Cp}\{\lambda_{f|r}(1-\varepsilon)\}] &= \frac{\sigma^2}{n}(1-\varepsilon)^{-1/(2q)}\ \mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\big|_{\lambda=\lambda_{f|r}} \\
&\times \left[o(1) + \{\sigma_u^2 \lambda_{f|r}(1-\varepsilon)n - \sigma^2\}\ \frac{\mathrm{tr}(\boldsymbol{S}_\lambda^2 - \boldsymbol{S}_\lambda^3)}{\sigma^2 \mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}\bigg|_{\lambda=\lambda_{f|r}}\right]\{1+o(1)\} \\
&= \left\{-\varepsilon\ \frac{\mathrm{tr}(\boldsymbol{S}_\lambda^2 - \boldsymbol{S}_\lambda^3)}{\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}\bigg|_{\lambda=\lambda_{f|r}} + o(1)\right\} \frac{\sigma^2 \mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n(1-\varepsilon)^{1/(2q)}}\bigg|_{\lambda=\lambda_{f|r}} < 0,
\end{aligned}$$

for $n \to \infty$, where $\sigma^2 = \sigma_u^2 \lambda_{f|r} n\{1+o(1)\}$ is used. Similarly, for $n \to \infty$

$$\mathrm{E}_\beta[T_{Cp}\{\lambda_{f|r}(1+\varepsilon)\}] = \left\{\varepsilon\ \frac{\mathrm{tr}(\boldsymbol{S}_\lambda^2 - \boldsymbol{S}_\lambda^3)}{\mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}\bigg|_{\lambda=\lambda_{f|r}} + o(1)\right\} \frac{\sigma^2 \mathrm{tr}(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)}{n(1+\varepsilon)^{1/(2q)}}\bigg|_{\lambda=\lambda_{f|r}} > 0.$$

Let now consider $T'_{Cp}(\tau\lambda_{f|r})/\mathrm{E}_\beta\{T'_{Cp}(\lambda_{f|r})\}$, where $\tau \in [1-\varepsilon, 1+\varepsilon]$, for any bounded $\varepsilon > 0$. It is easy to see that, since $\mathrm{var}_\beta\left\{T'_{Cp}(\tau\lambda_{f|r})\right\} = (\tau\lambda_{f|r})^{-2-1/(2q)} n^{-2} \text{const}\{1+o(1)\}$,

$$\mathrm{var}_\beta\left[\frac{T'_{Cp}(\tau\lambda_{f|r})}{\mathrm{E}_\beta\left\{T'_{Cp}(\lambda_{f|r})\right\}}\right] = O\left(\lambda_{f|r}^{1/(2q)}\right) \to 0, \quad n \to \infty.$$

Also,

$$\mathrm{E}_\beta\left\{T'_{Cp}(\tau\lambda_{f|r})\right\} = \mathrm{E}_\beta\left\{T'_{Cp}(\lambda_{f|r})\right\}\tau^{-1-1/(2q)}\left[1 + (\tau-1)\{1-1/(2q)\}\right]\{1+o(1)\},$$

where

$$\tau^{-1-1/(2q)}\left[1+(\tau-1)\{1-1/(2q)\}\right] = \begin{cases} 1+\varepsilon/q+O(\varepsilon^2), & \text{for } \tau = 1-\varepsilon \\ 1-\varepsilon/q+O(\varepsilon^2), & \text{for } \tau = 1+\varepsilon \end{cases},$$

so that for any fixed $\tau \in [1-\varepsilon, 1+\varepsilon]$ it holds $T'_{Cp}(\tau\lambda_{f|r})/\mathrm{E}_\beta\{T'_{Cp}(\lambda_{f|r})\} \xrightarrow{\mathcal{P}} 1$, as $n \to \infty$. Since $P(|\widetilde{\lambda}/\lambda_{f|r} - 1| \le \varepsilon) \to 1$ for $n \to \infty$ and any $\varepsilon > 0$ due to $\widehat{\lambda}_f \xrightarrow{\mathcal{P}} \lambda_{f|r}$, it follows

$$\frac{T'_{Cp}(\widetilde{\lambda})}{\mathrm{E}_\beta\left\{T'_{Cp}(\lambda_{f|r})\right\}} \xrightarrow{\mathcal{P}} 1.$$

Putting all together and applying Slutsky's lemma gives

$$\left(\frac{\widehat{\lambda}_f}{\lambda_{f|r}} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\lambda_{f|r}^{1/(2q)} c(\rho) \mathrm{sinc}\{\pi/(2q)\} \frac{4q(2q+1)}{3(2q-1)}\right).$$

3.2 Proof for $\widehat{\lambda}_r$

One needs to show

$$\frac{T_{ML}(\lambda_r) - \mathrm{E}_\beta\{T_{ML}(\lambda_r)\}}{\sqrt{\mathrm{var}_\beta\{T_{ML}(\lambda_r)\}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \text{and} \quad \frac{T'_{ML}(\widetilde{\lambda})}{\mathrm{E}_\beta\left\{T'_{ML}(\lambda_r)\right\}} \xrightarrow{\mathcal{P}} 1,$$

for some $\widetilde{\lambda}$ between $\widehat{\lambda}_r$ and λ_r. Simplifying Gamma functions in the expressions for $\mathrm{E}_\beta\left\{T'_{ML}(\lambda_r)\right\}$ and $\mathrm{var}_\beta\left\{T_{ML}(\lambda_r)\right\}$ obtained in Section 1.1.2 results in

$$\begin{aligned} \mathrm{E}_\beta\left\{T'_{ML}(\lambda_r)\right\} &= \frac{\sigma^2\lambda_r^{-1/(2q)-1}}{n\, c(\rho)} \frac{2q-1}{2q\,\mathrm{sinc}\{\pi/(2q)\}}\{1+o(1)\}, \\ \mathrm{var}_\beta\left\{T_{ML}(\lambda_r)\right\} &= \frac{2\sigma^4\lambda_r^{-1/(2q)}}{n^2\, c(\rho)} \frac{2q-1}{2q\,\mathrm{sinc}\{\pi/(2q)\}}\{1+o(1)\}. \end{aligned}$$

Consider now

$$\begin{aligned} n\,T_{ML}(\lambda_r) &= \left[\boldsymbol{Y}^t(\boldsymbol{S}_\lambda - \boldsymbol{S}_\lambda^2)\boldsymbol{Y} - \sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda)\right]\Big|_{\lambda=\lambda_{f|r}} + o_p(1), \\ \mathrm{E}_\beta\left\{n\,T_{ML}(\lambda_r)\right\} &= \sigma^2\mathrm{tr}(\boldsymbol{S}_\lambda)\Big|_{\lambda=\lambda_r} o(1). \end{aligned}$$

For $d_i = \sum_{j=1}^n \phi_{k,i}(x_j) y_j$, such that $\mathrm{E}_\beta(d_i^2) = \sigma^2 \lambda_r n \eta_i (1+\lambda_r n \eta_i)^{-1}\{1+o(1)\}$, let define random variables ξ_i by

$$\begin{aligned} n\ [T_{ML}(\lambda_r) - \mathrm{E}_f\{T_{ML}(\lambda_r)\}] &= \sum_{i=q+1}^{k+p+1} \left[d_i^2 - \sigma^2 \frac{1+\lambda_r n \eta_i}{\lambda_r n \eta_i}\{1+o(1)\}\right] \frac{(\lambda_r n \eta_i)}{(1+\lambda_r n \eta_i)^2} + o_p(1) \\ &=: \sum_{i=q+1}^{k+p+1} \xi_i, \end{aligned}$$

with $\mathrm{E}_\beta(\xi_i) = o(1)$ and $s_n^2 = \sum_{i=q+1}^{k+p+1} \mathrm{var}_\beta(\xi_i) = 2\sigma^4 \mathrm{tr}(\boldsymbol{S}_\lambda^2)\{1+o(1)\}$. Since $s_n^2 =$ const $\lambda_r^{-1/(2q)}$ and $(\lambda_r^{1/(2q)} k)^{-1} \to 0$ according to (A2), each $\mathrm{var}_\beta(\xi_i) = o(1)$ and there exist a constant B, such that $\mathrm{E}_\beta |\xi_i|^2 = \mathrm{var}_\beta(\xi_i) + o(1) < B$, $i = q+1, \ldots, k+p+1$. With this the Lyapunov's condition

$$s_n^{-4} \sum_{i=q+1}^{k+p+1} \mathrm{E}_\beta |\xi_i|^4 < B s_n^{-4} \sum_{i=q+1}^{k+p+1} \mathrm{E}_\beta |\xi_i|^2 = B s_n^{-2} = O\left(\lambda_r^{1/(2q)}\right)$$

converges to zero as n tends to infinity. Thus, $s_n^{-1} \sum_{i=q+1}^{k+p+1} \xi_i \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$, or equivalently, $[\mathrm{var}_\beta\{T_{ML}(\lambda_r)\}]^{-1/2} [T_{ML}(\lambda_r) - \mathrm{E}_\beta\{T_{ML}(\lambda_r)\}] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$.

Next is shown that $\widehat{\lambda}_r \xrightarrow{\mathcal{P}} \lambda_r$. From $\mathrm{var}_\beta\{T_{ML}(\lambda)\} = O\left(\lambda^{-1/(2q)} n^{-2}\right) \to 0$ for $n \to \infty$, it follows $T_{ML}(\lambda) \xrightarrow{\mathcal{P}} \mathrm{E}_\beta\{T_{ML}(\lambda)\}$. It remains to verify that $\mathrm{E}_\beta\left[T_{ML}\{\lambda_r(1-\varepsilon)\}\right] < 0 < \mathrm{E}_\beta\left[T_{ML}\{\lambda_r(1+\varepsilon)\}\right]$, for any $\varepsilon \in (0,1)$. Indeed,

$$\mathrm{E}_\beta[T_{ML}\{\lambda_r(1-\varepsilon)\}] = \left\{ -\varepsilon \left. \frac{\mathrm{tr}(\boldsymbol{S}_\lambda^2)}{\mathrm{tr}(\boldsymbol{S}_\lambda)} \right|_{\lambda=\lambda_r} + o(1) \right\} \left. \frac{\sigma^2 \mathrm{tr}(\boldsymbol{S}_\lambda)}{n(1-\varepsilon)^{1/(2q)}} \right|_{\lambda=\lambda_r} < 0,$$

for $n \to \infty$, where $\sigma^2 = \sigma_u^2 \lambda_r n\{1+o(1)\}$ is used. Similarly, for $n \to \infty$

$$\mathrm{E}_\beta[T_{ML}\{\lambda_r(1+\varepsilon)\}] = \left\{ \varepsilon \left. \frac{\mathrm{tr}(\boldsymbol{S}_\lambda^2)}{\mathrm{tr}(\boldsymbol{S}_\lambda)} \right|_{\lambda=\lambda_r} + o(1) \right\} \left. \frac{\sigma^2 \mathrm{tr}(\boldsymbol{S}_\lambda)}{n(1+\varepsilon)^{1/(2q)}} \right|_{\lambda=\lambda_r} > 0.$$

Let now consider $T_{ML}^{'}(\tau\lambda_r)/\mathrm{E}_\beta\{T_{ML}^{'}(\lambda_r)\}$, where $\tau \in [1-\varepsilon, 1+\varepsilon]$, for any bounded $\varepsilon > 0$. It is easy to see that, since $\mathrm{var}_\beta\left\{T_{ML}^{'}(\tau\lambda_r)\right\} = (\tau\lambda_r)^{-2-1/(2q)} n^{-2} \mathrm{const}\{1+o(1)\}$,

$$\mathrm{var}_\beta \left[\frac{T_{ML}^{'}(\tau\lambda_r)}{\mathrm{E}_\beta\left\{T_{ML}^{'}(\lambda_r)\right\}} \right] = O\left(\lambda_r^{1/(2q)}\right) \to 0, \quad n \to \infty.$$

Also,

$$\mathrm{E}_\beta\left\{T'_{ML}(\tau\lambda_r)\right\} \;=\; \mathrm{E}_\beta\left\{T'_{ML}(\lambda_r)\right\}\tau^{-1-1/(2q)}\left[1+(\tau-1)\{1-1/(2q)\}\right]\{1+o(1)\},$$

where

$$\tau^{-1-1/(2q)}\left[1+(\tau-1)\{1-1/(2q)\}\right] = \begin{cases} 1+\varepsilon/q+O(\varepsilon^2), & \text{for } \tau=1-\varepsilon \\ 1-\varepsilon/q+O(\varepsilon^2), & \text{for } \tau=1+\varepsilon \end{cases},$$

so that for any fixed $\tau \in [1-\varepsilon, 1+\varepsilon]$ it holds $T'_{ML}(\tau\lambda_r)/\mathrm{E}_\beta\{T'_{ML}(\lambda_r)\} \xrightarrow{\mathcal{P}} 1$, as $n \to \infty$. Since $P(|\widetilde{\lambda}/\lambda_r - 1| \le \varepsilon) \to 1$ for $n \to \infty$ and any $\varepsilon > 0$ due to $\widehat{\lambda}_r \xrightarrow{\mathcal{P}} \lambda_r$, it follows

$$\frac{T'_{ML}(\widetilde{\lambda})}{\mathrm{E}_\beta\left\{T'_{ML}(\lambda_r)\right\}} \xrightarrow{\mathcal{P}} 1.$$

Putting all together and applying Slutsky's lemma gives

$$\left(\frac{\widehat{\lambda}_r}{\lambda_r} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\lambda_r^{1/(2q)} c(\rho)\mathrm{sinc}\{\pi/(2q)\}\frac{2q}{2q-1}\right).$$

4 Data-driven selection of q

Using the same functions f_1 and f_2 and the same setting as in Section 4 of the paper, $R^*(q)$ was calculated for $q = 2, 3, 4, 5$ and two sample sizes $n = 350$ and $n = 1000$, fixing the number of knots at $k = 40$. The results from 500 Monte Carlo replications are shown in Figure 1 and agree with the simulation results from Section 4. For f_1 using $q = 3$ or $q = 4$ for $n = 350$ and $q = 4$ for $n = 1000$ seem to do best, since the corresponding $|R^*(q)|$ is smallest. For f_2 using $q = 4$ is more advisable.

References

Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institue of Statistical Mathematics*, 21:243 – 47.

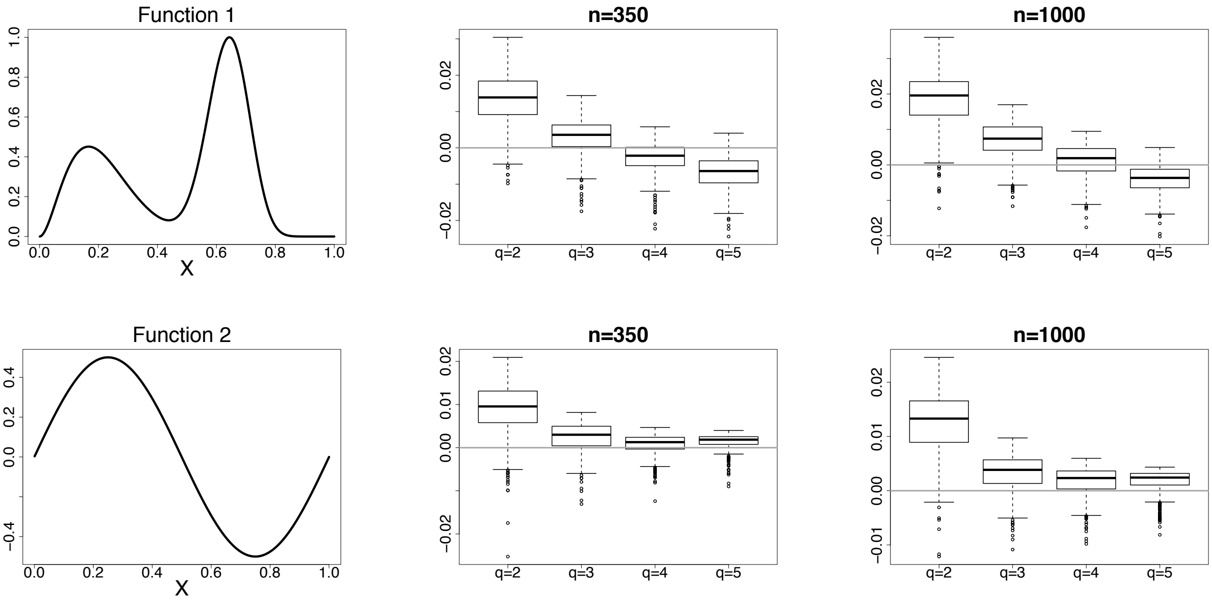

Figure 1: Choice of the optimal q: Boxplots of $R^*(q)$ for different values of q for $n = 350$ (middle plots) and $n = 1000$ (right plots) for f_1 (top left) and f_2 (bottom left).

Claeskens, G., Krivobokova, T., and Opsomer, J. (2009). Asymptotic properties of penalized spline estimators. *Biometrika*, 96(6):529–544.

Craven, P. and Wahba, G. (1978). Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numer. Math.*, 31(4):377–403.

Patterson, H. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545–554.

van der Vaart, A. W. (1998). *Asymptotic statistics.* Cambridge University Press, New York.

Wiens, D. P. (1992). On moments of quadratic forms in non-spherically distributed variables. *Statistics*, 23:265–270.