Supplement to "Consistency and Asymptotic Normality of Sieve ML Estimators Under Low-Level Conditions"

Herman J. Bierens
Pennsylvania State University

1. Introduction

This online supplement to Bierens (2014) contains the omitted proofs. Throughout I will use the same notations as in Bierens (2014), as follows. The indicator function is denoted by $I(.)$, and $\mathbb{N}$ and $\mathbb{N}_0$ denote the sets of positive and non-negative integers, respectively. The partial derivative to a parameter with index k will be denoted by ∇_k, and $\nabla_{k,m}$ denotes the second partial derivatives to parameters with indices k and m. To distinguish infinite dimensional parameters from finite dimensional ones, the former are displayed in bold face. Following Billingsley (1968), I will use the double-arrow "$\Rightarrow$" to indicate weak convergence of sequences of random function in the metric space $\mathcal{C}[0,1]$ of continuous real functions on $[0,1]$, endowed with the metric $\sup_{0\leq u\leq 1}|f(u)-g(u)|$, and following van der Vaart (1998), the wiggling arrow "$\rightsquigarrow$" indicates weak convergence of a sequence of random elements in a Hilbert space. Finally, the operator π_n applied to an infinite sequence $\boldsymbol{\delta}=\{\delta_m\}_{m=1}^{\infty}$ replaces all the δ_m's for $m>n$ by zeros.

2. Proof of Lemma 2.1

Consider first the case that X is a single random variable, and $\beta_0\neq 0$, and suppose that there exist a distribution function H on $[0,1]$ and coefficients α_*, β_* such that $H_0(G(\alpha_0+\beta_0 X))=H(G(\alpha_*+\beta_* X))$ a.s. Obviously, this is only possible if $\beta_*\neq 0$, with the same sign as β_0. Next, denote $Z=\alpha_*+\beta_* X$ and suppose that the distribution of Z has support $\mathbb{R}$. Then

$$H(G(z))=H_0(G(\alpha_0-c\alpha_*+c.z)) \text{ for all } z\in\mathbb{R},$$

where $c = \beta_0/\beta_* > 0$. For $0 < u_1 < u_2 < 1$, let $z_1 = G^{-1}(u_1),\ z_2 = G^{-1}(u_2)$. Then under the quantile conditions $H(u_1) = H_0(u_1) = u_1,\ H(u_2) = H_0(u_2) = u_2$,

$$\begin{aligned} H(u_1) &= H_0(G(\alpha_0 - c\alpha_* + c.z_1)) = H_0(u_1) = H_0(G(z_1)), \\ H(u_2) &= H_0(G(\alpha_0 - c\alpha_* + c.z_2)) = H_0(u_2) = H_0(G(z_2)), \end{aligned}$$

hence by the strict monotonicity of $H_0(G(z))$,

$$\alpha_0 - c\alpha_* + c.z_1 = z_1,\ \alpha_0 - c\alpha_* + c.z_2 = z_2,$$

which implies $c = \beta_0/\beta_* = 1$ and $\alpha_0 = \alpha_*$.

Consider now the case $X = (X_1, X_2')' \in \mathbb{R}^q$, $q \geq 2$, where $X_2 \in \mathbb{R}^{q-1}$. Let us assume again that there exist an absolutely conditional distribution function H and parameters α_* and $\beta_* = \left(\beta_{*,1}, \beta_{*,2}'\right)' \in \mathbb{R}^{p-1}$ such that

$$H_0\left(G(\alpha_0 + \beta_{0,1}X_1 + \beta_{0,2}'X_2)\right) = H\left(G(\alpha_* + \beta_{*,1}X_1 + \beta_{*,2}'X_2)\right) \text{ a.s.}$$

Moreover, suppose that the conditional distribution of X_1 given X_2 has support $\mathbb{R}$, and that $\beta_{0,1} \neq 0$. It follows from the previous argument and the quantile restrictions that conditional on X_2, $\beta_{*,1} = \beta_{0,1}$ and

$$\alpha_* + \beta_{*,2}'X_2 = \alpha_0 + \beta_{0,2}'X_2 \text{ a.s.} \tag{2.1}$$

Assuming that $E[X_2'X_2] < \infty$, so that $E[X_2]$ is defined and $\text{Var}(X_2)$ is finite, (2.1) implies that $(\beta_{*,2} - \beta_{0,2})'\left(X_2 - E\left[X_2\right]\right) = 0$ a.s., hence

$$(\beta_{*,2} - \beta_{0,2})'\text{Var}\left(X_2\right)\left(\beta_{*,2} - \beta_{0,2}\right) = 0$$

Therefore, if $\text{Var}(X_2)$ is nonsingular then $\beta_{*,2} = \beta_{0,2}$ so that by (2.1), $\alpha_* = \alpha_0$. The lemma follows now from the fact that the nonsingularity of $\text{Var}(X_2)$ is implied by the nonsingularity of $\text{Var}(X)$.

3. Proof of Lemma 4.1

It follows from the mean value theorem and the choice of G as the logistic distribution function that

$$\begin{aligned} |G((1,X')\theta_1) - G((1,X')\theta_2)| &\leq |(1,X')(\theta_1 - \theta_2)| \sup_x G(x)(1 - G(x)) \\ &\leq ||\theta_1 - \theta_2||.(1 + ||X||)/4 \end{aligned}$$

whereas for $\boldsymbol{\delta}_i = \{\delta_{i,m}\}_{m=1}^{\infty}$, $i = 1, 2$,

$$\sup_{0\leq u\leq 1} |H(u|\boldsymbol{\delta}_1) - H(u|\boldsymbol{\delta}_2)| \leq \int_0^1 |h(u|\boldsymbol{\delta}_1) - h(u|\boldsymbol{\delta}_2)|\,\mathrm{d}u = O\left(\sqrt{\sum_{m=1}^{\infty}(\delta_{1,m} - \delta_{2,m})^2}\right)$$

similar to Theorem 3.1. Hence

$$\begin{aligned}
&|H(G((1,X')\theta_1)|\boldsymbol{\delta}_1) - H(G((1,X')\theta_2)|\boldsymbol{\delta}_2)| \\
&\quad \leq |H(G((1,X')\theta_1)|\boldsymbol{\delta}_1) - H(G((1,X')\theta_2)|\boldsymbol{\delta}_1)| \\
&\qquad + |H(G((1,X')\theta_2)|\boldsymbol{\delta}_1) - H(G((1,X')\theta_2)|\boldsymbol{\delta}_2)| \\
&\quad \leq H\left(||\theta_1 - \theta_2||.(1 + ||X||)/4|\,\boldsymbol{\delta}_1\right) + O\left(\sqrt{\sum_{m=1}^{\infty}(\delta_{1,m} - \delta_{2,m})^2}\right)
\end{aligned}$$

The lemma under review follows now from the continuity of $H\left(u|\boldsymbol{\delta}_1\right)$ in u.

4. Proof of Theorem 4.1

Let $\alpha > 0$ be arbitrary. Then by condition (i) there exists an $n_0(\alpha) \in \mathbb{N}$ such that $Q(\boldsymbol{\xi}_{n_0(\alpha)}) \geq Q(\boldsymbol{\xi}^0) - \alpha$, hence for $n_N \geq n_0(\alpha)$,

$$\begin{aligned}
\widehat{Q}_N(\widehat{\boldsymbol{\xi}}_N) &\geq \widehat{Q}_N(\boldsymbol{\xi}_{n_0(\alpha)}) \geq \widehat{Q}_N(\boldsymbol{\xi}_{n_0(\alpha)}) - Q(\boldsymbol{\xi}_{n_0(\alpha)}) + Q(\boldsymbol{\xi}^0) - \alpha \\
&\geq Q(\boldsymbol{\xi}^0) - \alpha - R_N(\alpha), \qquad (4.1)
\end{aligned}$$

where

$$R_N(\alpha) = \left|\widehat{Q}_N(\boldsymbol{\xi}_{n_0(\alpha)}) - Q(\boldsymbol{\xi}_{n_0(\alpha)})\right| \overset{\text{a.s.}}{\to} 0.$$

The latter follows from Kolmogorov's strong law of large numbers.

Denote $\Xi_c(\varepsilon) = \{\boldsymbol{\xi} \in \Xi_c : d(\boldsymbol{\xi}, \boldsymbol{\xi}^0) \geq \varepsilon\}$, so that the result of Theorem 4.1 reads:

$$\lim_{N\to\infty} \Pr[\widehat{\boldsymbol{\xi}}_N \in \Xi_c(\varepsilon)] = 0.$$

If $\Xi_c(\varepsilon) = \emptyset$ then trivially, $\Pr[\widehat{\boldsymbol{\xi}}_N \in \Xi_c(\varepsilon)] = 0$. Therefore, assume that $\Xi_c(\varepsilon) \neq \emptyset$. Then $\Xi_c(\varepsilon)$ is compact.[1]

[1] See for example the proof of Theorem II.6 on page 290 in Bierens (2004).

Next, observe that for each $\boldsymbol{\xi} \in \Xi_c(\varepsilon)$, $\sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < \eta} f(Z, \boldsymbol{\xi}_*) - f(Z, \boldsymbol{\xi})$ is a.s. non-negative and monotonic increasing in $\eta \in (0, \infty)$, hence if $E[f(Z, \boldsymbol{\xi})] > -\infty$ then by condition (d) and (e) and the dominated convergence theorem,

$$\begin{aligned} \lim_{\eta \downarrow 0} E &\left[\sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < \eta} f(Z, \boldsymbol{\xi}_*) - f(Z, \boldsymbol{\xi}) \right] \\ &= E \left[\lim_{\eta \downarrow 0} \sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < \eta} f(Z, \boldsymbol{\xi}_*) - f(Z, \boldsymbol{\xi}) \right] \\ &= E[f(Z, \boldsymbol{\xi}) - f(Z, \boldsymbol{\xi})] = 0. \end{aligned}$$

Thus, for each $\boldsymbol{\xi} \in \Xi_c(\varepsilon)$ for which $E[f(Z, \boldsymbol{\xi})] > -\infty$ we have

$$\begin{aligned} \lim_{\eta \downarrow 0} E \left[\sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < \eta} f(Z, \boldsymbol{\xi}_*) \right] &= E[f(Z, \boldsymbol{\xi})] \le \sup_{\boldsymbol{\xi} \in \Xi_c(\varepsilon)} Q(\boldsymbol{\xi}) \\ &= Q(\boldsymbol{\xi}^0) - \left(Q(\boldsymbol{\xi}^0) - \sup_{\boldsymbol{\xi} \in \Xi_c(\varepsilon)} Q(\boldsymbol{\xi}) \right). \end{aligned}$$

Now choose $\alpha > 0$ such that $Q(\boldsymbol{\xi}^0) - \sup_{\boldsymbol{\xi} \in \Xi_c(\varepsilon)} Q(\boldsymbol{\xi}) > 2\alpha$. Then there exists an $\eta(\boldsymbol{\xi}, \alpha) > 0$ such that

$$E \left[\sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < \eta(\boldsymbol{\xi}, \alpha)} f(Z, \boldsymbol{\xi}_*) \right] < Q(\boldsymbol{\xi}^0) - 2\alpha. \tag{4.2}$$

In the case $E[f(Z, \boldsymbol{\xi})] = -\infty$ it follows from condition (e) and the monotone convergence theorem that

$$\begin{aligned} \lim_{\eta \downarrow 0} E &\left[\sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < \eta} f(Z, \boldsymbol{\xi}_*) \right] \\ &= E[\overline{f}(Z)] - \lim_{m \to \infty} E \left[\overline{f}(Z) - \sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < 1/m} f(Z, \boldsymbol{\xi}_*) \right] \\ &= E[\overline{f}(Z)] - E \left[\overline{f}(Z) - \lim_{m \to \infty} \sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < 1/m} f(Z, \boldsymbol{\xi}_*) \right] \\ &= E[f(Z, \boldsymbol{\xi})] = -\infty, \end{aligned}$$

so that also in this case there exists an $\eta(\boldsymbol{\xi},\alpha) > 0$ such that (4.2) holds.

By the compactness of $\Xi_c(\varepsilon)$ there exist a finite number of elements $\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_K$ of $\Xi_c(\varepsilon)$ such that

$$\Xi_c(\varepsilon) \subset \cup_{i=1}^K \{\boldsymbol{\xi}_* \in \Xi : d(\boldsymbol{\xi}_*, \boldsymbol{\xi}_i) < \eta(\boldsymbol{\xi}_i, \alpha)\}.$$

Hence, denoting

$$g(z, \boldsymbol{\xi}|\alpha) = \sup_{\boldsymbol{\xi}_* \in \Xi:\ d(\boldsymbol{\xi}_*, \boldsymbol{\xi}) < \eta(\boldsymbol{\xi}, \alpha)} f(z, \boldsymbol{\xi}_*),$$

it follows from (4.2) that

$$\begin{aligned}
\sup_{\boldsymbol{\xi} \in \Xi_c(\varepsilon)} \widehat{Q}_N(\boldsymbol{\xi}) &\leq \max_{1 \leq i \leq K} \left(\frac{1}{N} \sum_{j=1}^N g(Z_j, \boldsymbol{\xi}_i|\alpha) \right) \\
&\leq \max_{1 \leq i \leq K} \left(\frac{1}{N} \sum_{j=1}^N g(Z_j, \boldsymbol{\xi}_i|\alpha) - E[g(Z, \boldsymbol{\xi}_i|\alpha)] \right) \\
&\quad + \max_{1 \leq i \leq K} \left(E[g(Z, \boldsymbol{\xi}_i|\alpha)] \right) \\
&< S_N(\alpha) + Q(\boldsymbol{\xi}^0) - 2\alpha, \qquad (4.3)
\end{aligned}$$

where

$$S_N(\alpha) = \max_{1 \leq i \leq K} \left| \frac{1}{N} \sum_{j=1}^N g(Z_j, \boldsymbol{\xi}_i|\alpha) - E[g(Z, \boldsymbol{\xi}_i|\alpha)] \right| \overset{\text{a.s.}}{\to} 0.$$

Note that by conditions (e) and (j), $E[|g(Z, \boldsymbol{\xi}_i|\alpha)|] < \infty$ for $i = 1, 2, .., K$, so that the convergence result involved follows from Kolmogorov's strong law of large numbers.

Combining (4.1) and (4.3), it follows now that $\widehat{\boldsymbol{\xi}}_N \in \Xi_c(\varepsilon)$ implies $Q(\boldsymbol{\xi}^0) - \alpha - R_N(\alpha) \leq \widehat{Q}_N(\widehat{\boldsymbol{\xi}}_N) < S_N(\alpha) + Q(\boldsymbol{\xi}^0) - 2\alpha$ and thus $S_N(\alpha) + R_N(\alpha) \geq \alpha$, so that

$$\Pr[\widehat{\boldsymbol{\xi}}_N \in \Xi_c(\varepsilon)] \leq \Pr[S_N(\alpha) + R_N(\alpha) \geq \alpha] \to 0.$$

Finally, let Ξ_c be an arbitrary compact subset of Ξ containing $\boldsymbol{\xi}^0$ in its interior. Then $\inf_{\boldsymbol{\xi} \in \Xi \backslash \Xi_c} d(\boldsymbol{\xi}, \boldsymbol{\xi}^0) = \varepsilon > 0$, hence $\Pr[\widehat{\boldsymbol{\xi}}_N \in \Xi \backslash \Xi_c] \leq \Pr[d(\widehat{\boldsymbol{\xi}}_N, \boldsymbol{\xi}^0) \geq \varepsilon] \to 0$ if $\text{plim}_{N\to\infty} d(\widehat{\boldsymbol{\xi}}_N, \boldsymbol{\xi}^0) = 0$, so that the latter implies $\lim_{N\to\infty} \Pr[\widehat{\boldsymbol{\xi}}_N \in \Xi_c] = 1$. The other way around, i.e., $\lim_{N\to\infty} \Pr[\widehat{\boldsymbol{\xi}}_N \in \Xi_c] = 1$ implies $\text{plim}_{N\to\infty} d(\widehat{\boldsymbol{\xi}}_N, \boldsymbol{\xi}^0) = 0$, follows from the first part of Theorem 4.1.

5. Proof of Theorem 4.2

The case in Assumption 4.2(a) follows trivially from Theorem 4.1. In the case of Assumption 4.2(b), denote

$$\alpha = \frac{1}{2}\left(Q(\boldsymbol{\xi}^0) - E\left[\sup_{\boldsymbol{\xi}\in\overline{\Xi}\backslash\Xi_c} f(Z,\boldsymbol{\xi})\right]\right).$$

Then

$$\sup_{\boldsymbol{\xi}\in\overline{\Xi}\backslash\Xi_c} \widehat{Q}_N(\boldsymbol{\xi}) \leq \frac{1}{N}\sum_{j=1}^{N}\sup_{\boldsymbol{\xi}\in\overline{\Xi}\backslash\Xi_c} f(Z_j,\boldsymbol{\xi}) \leq T_N + Q(\boldsymbol{\xi}^0) - 2\alpha,$$

where

$$T_N = \left|\frac{1}{N}\sum_{j=1}^{N}\sup_{\boldsymbol{\xi}\in\overline{\Xi}\backslash\Xi_c} f(Z_j,\boldsymbol{\xi}) - E\left[\sup_{\boldsymbol{\xi}\in\overline{\Xi}\backslash\Xi_c} f(Z,\boldsymbol{\xi})\right]\right| \stackrel{\text{a.s.}}{\to} 0.$$

It follows now similar to the proof of Theorem 4.1 that $\widehat{\boldsymbol{\xi}}_N \in \overline{\Xi}\backslash\Xi_c$ implies $R_N(\alpha)+T_N \geq \alpha$, hence $\lim_{N\to\infty}\Pr[\widehat{\boldsymbol{\xi}}_N \in \overline{\Xi}\backslash\Xi_c] = 0$ and thus $\lim_{N\to\infty}\Pr[\widehat{\boldsymbol{\xi}}_N \in \Xi_c] = 1$. Combined with Theorem 4.1, the latter implies that $\lim_{N\to\infty}\Pr[d(\widehat{\boldsymbol{\xi}}_N,\boldsymbol{\xi}^0) \geq \varepsilon] = 0$.

6. Proof of Lemma 5.1

I will prove Lemma 5.1 via the following sub-lemmas. In each sub-lemma, $m = 0, 1, ..., \ell$, and for each m, $C_m \in (0,\infty)$ is a generic constant depending on m and the norm $||\boldsymbol{\delta}||_m$.

Lemma 5.1(a). $\sup_{0\leq u\leq 1}|h^{(m)}(u|\boldsymbol{\delta})| \leq C_m$

Proof. Write

$$h(u|\boldsymbol{\delta}) = \frac{f(u|\boldsymbol{\delta})^2}{1+\sum_{m=1}^{\infty}\delta_m^2}$$

where

$$f(u|\boldsymbol{\delta}) = 1 + \sqrt{2}\sum_{k=1}^{\infty}\delta_k\cos(k\pi u)$$

and denote $f^{(0)}(u|\boldsymbol{\delta}) = f(u|\boldsymbol{\delta})$, $f^{(m)}(u|\boldsymbol{\delta}) = \mathrm{d}^m f(u|\boldsymbol{\delta})/(\mathrm{d}u)^m$ for $m = 1, ..., \ell$. Observe that for $m = 0, 1, ..., \ell$

$$\begin{aligned} f^{(m)}(u|\boldsymbol{\delta}) &= I(m=0) + \sqrt{2}\pi^m \sum_{k=1}^{\infty} k^m \delta_k c_m(k\pi u), \qquad (6.1) \\ & \text{where } c_m(x) = \mathrm{d}^m \cos(x)/(\mathrm{d}x)^m, \end{aligned}$$

hence

$$\sup_{0\leq u\leq 1} |f^{(m)}(u|\boldsymbol{\delta})| \leq \max(1, \sqrt{2}\pi^m ||\boldsymbol{\delta}||_m) \qquad (6.2)$$

Moreover,

$$\begin{aligned} h^{(1)}(u|\boldsymbol{\delta}) &= \frac{2f(u|\boldsymbol{\delta})f^{(1)}(u|\boldsymbol{\delta})}{1+\sum_{k=1}^{\infty}\delta_k^2} \\ h^{(2)}(u|\boldsymbol{\delta}) &= \frac{2f^{(1)}(u|\boldsymbol{\delta})^2 + 2f(u|\boldsymbol{\delta})f^{(2)}(u|\boldsymbol{\delta})}{1+\sum_{k=1}^{\infty}\delta_k^2}, \\ h^{(3)}(u|\boldsymbol{\delta}) &= \frac{2f(u|\boldsymbol{\delta})f^{(3)}(u|\boldsymbol{\delta}) + 6f^{(1)}(u|\boldsymbol{\delta})f^{(2)}(u|\boldsymbol{\delta})}{1+\sum_{k=1}^{\infty}\delta_k^2} \end{aligned}$$

etcetera, and more generally,

$$h^{(m)}(u|\boldsymbol{\delta}) = \frac{\sum_{k=0}^{m} \omega_{k,m} f^{(m-k)}(u|\boldsymbol{\delta})f^{(k)}(u|\boldsymbol{\delta})}{1+\sum_{k=1}^{\infty}\delta_k^2} \qquad (6.3)$$

for $m = 0, 1, 2, ..., \ell$, where $\omega_{k,m} \in \mathbb{N}_0$. Since obviously $||\boldsymbol{\delta}||_k \leq ||\boldsymbol{\delta}||_m$ for $0 \leq k \leq m$ it follows therefore from (6.2) and (6.3) that

$$\begin{aligned} \sup_{0\leq u\leq 1} |h^{(m)}(u|\boldsymbol{\delta})| &\leq \sum_{k=0}^{m} \omega_{k,m} \sup_{0\leq u\leq 1} |f^{(m-k)}(u|\boldsymbol{\delta})| \sup_{0\leq u\leq 1} |f^{(k)}(u|\boldsymbol{\delta})| \\ &\leq \left(\sum_{k=0}^{m} \omega_{k,m}\right) \max\left(1, 2\pi^{2m}(||\boldsymbol{\delta}||_m)^2\right) \\ &\leq \left(\sum_{k=0}^{m} \omega_{k,m}\right) \left(1 + 2\pi^{2m}(||\boldsymbol{\delta}||_m)^2\right), \end{aligned}$$

which proves Lemma 5.1(a). ■

Lemma 5.1(b). $\sup_{0\leq u\leq 1}|\nabla_i h^{(m)}(u|\boldsymbol{\delta})| < C_m.i^m$.

Proof. It follows from (6.1) that for $i\in\mathbb{N}$,

$$\nabla_i f^{(m)}(u|\boldsymbol{\delta}) = \frac{\partial f^{(m)}(u|\boldsymbol{\delta})}{\partial \delta_i} = \sqrt{2}(\pi.i)^m c_m(i.\pi u). \tag{6.4}$$

Moreover, it follows from (6.3) that

$$\begin{aligned}\nabla_i\left(\left(1+\sum_{k=1}^{\infty}\delta_k^2\right)h^{(m)}(u|\boldsymbol{\delta})\right) &= 2\delta_i h^{(m)}(u|\boldsymbol{\delta}) + \left(1+\sum_{k=1}^{\infty}\delta_k^2\right)\nabla_i h^{(m)}(u|\boldsymbol{\delta}) \\ &= \sqrt{2}\sum_{k=0}^{m}\omega_{k,m}(\pi.i)^{m-k}c_{m-k}(i.\pi u)f^{(k)}(u|\boldsymbol{\delta}) \\ &\quad +\sqrt{2}\sum_{k=0}^{m}\omega_{k,m}(\pi.i)^{k}c_{k}(i.\pi u)f^{(m-k)}(u|\boldsymbol{\delta}),\end{aligned}$$

so that

$$\begin{aligned}\nabla_i h^{(m)}(u|\boldsymbol{\delta}) = \frac{1}{1+\sum_{k=1}^{\infty}\delta_k^2}&\left\{\sqrt{2}\sum_{k=0}^{m}\omega_{k,m}(\pi.i)^{m-k}c_{m-k}(i.\pi u)f^{(k)}(u|\boldsymbol{\delta})\right. \\ &\left. +\sqrt{2}\sum_{k=0}^{m}\omega_{k,m}(\pi.i)^{k}c_{k}(i.\pi u)f^{(m-k)}(u|\boldsymbol{\delta}) - 2\delta_i h^{(m)}(u|\boldsymbol{\delta})\right\}\end{aligned} \tag{6.5}$$

hence

$$\begin{aligned}\left|\nabla_i h^{(m)}(u|\boldsymbol{\delta})\right| &\leq \sqrt{2}i^m\pi^m\sum_{k=0}^{m}\omega_{k,m}\left(|f^{(k)}(u|\boldsymbol{\delta})| + |f^{(m-k)}(u|\boldsymbol{\delta})|\right) \\ &\quad +2|\delta_i|.|h^{(m)}(u|\boldsymbol{\delta})| \\ &\leq \sqrt{2}i^m\pi^m\sum_{k=0}^{m}\omega_{k,m}\left(|f^{(k)}(u|\boldsymbol{\delta})| + |f^{(m-k)}(u|\boldsymbol{\delta})|\right) \\ &\quad +2i^m||\boldsymbol{\delta}||_m.|h^{(m)}(u|\boldsymbol{\delta})|\end{aligned}$$

The result involved now follows straightforwardly from Lemma 5.1(a), (6.2), (6.4) and (6.5). ■

Lemma 5.1(c). $\sup_{0\leq u\leq 1}\left|\nabla_{i_1,i_2}h^{(m)}(u|\boldsymbol{\delta})\right| \leq C_m.i_1^m i_2^m$.

Proof. It follows straightforwardly from (6.5) that

$$\nabla_{i_1,i_2}h^{(m)}(u|\boldsymbol{\delta}) = \frac{1}{1+\sum_{k=1}^{\infty}\delta_k^2}\left\{2\pi^m\sum_{k=0}^{m}\omega_{k,m}i_1^{m-k}i_2^k c_{m-k}(i_1.\pi u)c_k(i_2.\pi u)\right.$$
$$+ 2\pi^m\sum_{k=0}^{m}\omega_{k,m}i_2^{m-k}i_1^k c_{m-k}(i_2.\pi u)c_k(i_1.\pi u)$$
$$\left. - 2\delta_{i_1}\nabla_{i_2}h^{(m)}(u|\boldsymbol{\delta}) - 2\delta_{i_2}\nabla_{i_1}h^{(m)}(u|\boldsymbol{\delta}) - 2.I(i_1=i_2)h^{(m)}(u|\boldsymbol{\delta})\right\} \qquad (6.6)$$

It follows therefore similar to the proof of Lemma 5.1(b) that the result of Lemma 5.1(c) holds. ■

Lemma 5.1(d). $\sup_{0\leq u\leq 1}\left|h^{(m)}(u|\boldsymbol{\delta}_*) - h^{(m)}(u|\boldsymbol{\delta})\right| \leq C_m(||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_m+||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_m^2)$.

Proof. It follows trivially from (6.1) that

$$\sup_{0\leq u\leq 1}|f^{(m)}(u|\boldsymbol{\delta}_*) - f^{(m)}(u|\boldsymbol{\delta})| \leq \sqrt{2}\pi^m||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_m$$

and therefore by (6.1),

$$\sup_{0\leq u\leq 1}\left|f^{(m-k)}(u|\boldsymbol{\delta}_*)f^{(k)}(u|\boldsymbol{\delta}_*) - f^{(m-k)}(u|\boldsymbol{\delta})f^{(k)}(u|\boldsymbol{\delta})\right|$$
$$\leq \sup_{0\leq u\leq 1}|f^{(m-k)}(u|\boldsymbol{\delta}_*) - f^{(m-k)}(u|\boldsymbol{\delta})|.\sup_{0\leq u\leq 1}|f^{(k)}(u|\boldsymbol{\delta}_*) - f^{(k)}(u|\boldsymbol{\delta})|$$
$$+ \sup_{0\leq u\leq 1}|f^{(k)}(u|\boldsymbol{\delta})|\sup_{0\leq u\leq 1}|f^{(m-k)}(u|\boldsymbol{\delta}_*) - f^{(m-k)}(u|\boldsymbol{\delta})|$$
$$+ \sup_{0\leq u\leq 1}|f^{(m-k)}(u|\boldsymbol{\delta})|\sup_{0\leq u\leq 1}|f^{(k)}(u|\boldsymbol{\delta}_*) - f^{(k)}(u|\boldsymbol{\delta})|$$
$$\leq 2\pi^m||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_{m-k}.||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_k$$
$$+ \sqrt{2}\pi^{m-k}\max(1,\sqrt{2}\pi^k||\boldsymbol{\delta}||_k)||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_{m-k}$$
$$+ \sqrt{2}\pi^k\max(1,\sqrt{2}\pi^{m-k}||\boldsymbol{\delta}||_{m-k})||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_k$$
$$\leq 2\pi^m\left(||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_m\right)^2 + 4\pi^m\max(1,||\boldsymbol{\delta}||_m)||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_m$$
$$\leq 4\pi^m\max(1,||\boldsymbol{\delta}||_m).(||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_m + ||\boldsymbol{\delta}_*-\boldsymbol{\delta}||_m^2).$$

Moreover, it follows from (6.3) that

$$\left|h^{(m)}(u|\boldsymbol{\delta}_*) - h^{(m)}(u|\boldsymbol{\delta})\right|$$

$$
\leq \sum_{k=0}^{m} \omega_{k,m} \sup_{0\leq u\leq 1} \left| f^{(m-k)}(u|\boldsymbol{\delta}_*) f^{(k)}(u|\boldsymbol{\delta}_*) - f^{(m-k)}(u|\boldsymbol{\delta}) f^{(k)}(u|\boldsymbol{\delta}) \right|
$$

$$
+ \sum_{k=1}^{\infty} \left|\delta_{*,k}^2 - \delta_k^2\right| . \sup_{0\leq u\leq 1} |h^{(m)}(u|\boldsymbol{\delta})|
$$

Furthermore,

$$
\begin{aligned}
\sum_{k=1}^{\infty} \left|\delta_{*,k}^2 - \delta_k^2\right| &\leq \sum_{k=1}^{\infty} \left(|\delta_{*,k} - \delta_k| + 2|\delta_k|\right) . |\delta_{2,k} - \delta_{1,k}| \\
&\leq \left(\sum_{k=1}^{\infty} |\delta_{*,k} - \delta_k|\right)^2 + 2\left(\sum_{k=1}^{\infty} |\delta_k|\right)\left(\sum_{k=1}^{\infty} |\delta_{*,k} - \delta_k|\right) \\
&= ||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_0^2 + 2||\boldsymbol{\delta}||_0 . ||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_0 \\
&\leq ||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_m^2 + 2||\boldsymbol{\delta}||_m . ||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_m . \qquad (6.7)
\end{aligned}
$$

These inequalities together with the result of Lemma 5.1(a) prove Lemma 5.1(d). ■

Lemma 5.1(e). $\sup_{0\leq u\leq 1} \left|\nabla_i h^{(m)}(u|\boldsymbol{\delta}_1) - \nabla_i h^{(m)}(u|\boldsymbol{\delta}_2)\right| < C_m . i^m . (||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_m + ||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_m^2)$

Proof. This result follows straightforwardly from (6.5) and a similar argument as in the proof of Lemma 5.1(d). ■

Lemma 5.1(f). $\sup_{0\leq u\leq 1} \left|\nabla_{i_1,i_2} h^{(m)}(u|\boldsymbol{\delta}_*) - \nabla_{i_1,i_2} h^{(m)}(u|\boldsymbol{\delta})\right| < C_m . i_1^m i_2^m . (||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_m + ||\boldsymbol{\delta}_* - \boldsymbol{\delta}||_m^2)$.

Proof. Using (6.6), this result follows similar to the proof of Lemma 5.1(d). ■

7. Proof of Lemma 6.1

Recall that by the first-order condition it follows that for $k \leq n$, $\sum_{j=1}^{N} \nabla_k f_j(\widehat{\boldsymbol{\xi}}_n) = 0$ whenever $\widehat{\boldsymbol{\xi}}_n \in \Xi_n^{\text{Int}}$. Moreover, it follows from Assumptions 6.1(c) and 6.2 that for any fixed $K \leq n$, and $n \to \infty$,

$$
\Pr[(\widehat{\xi}_{n,1}, ..., \widehat{\xi}_{n,K})' \in \Xi_K^{\text{Int}}] = P_n(K) \to 1.
$$

This implies that there exists a sequence K_n converging to infinity with n such that also $\lim_{n\to\infty} P_n(K_n) \to 1$. To see this, denote $\underline{P}_n(K) = \inf_{m\geq n} P_m(K)$ and note that for each K, $\underline{P}_n(K) \uparrow 1$ monotonically as $n \to \infty$, and for each n, $\underline{P}_n(K)$ is non-increasing in K. For $m \geq 1$, let n_m be the smallest $n \geq m$ for which $\underline{P}_n(m) + m^{-2} \geq 1$, for example. Obviously, $n_1 = 1$. Moreover, since $\underline{P}_n(m) + m^{-2} > \underline{P}_n(m+1) + (m+1)^{-2}$, we have $n_{m+1} > n_m$. Now let $K_{m,n} = m$ for $n_m \leq n \leq n_{m+1} - 1$. Then for all $m \in \mathbb{N}$,

$$\Pr[(\widehat{\xi}_{n,1}, ..., \widehat{\xi}_{n,K_{m,n}})' \in \Xi^{\text{Int}}_{K_{m,n}}] \geq 1 - m^{-2}.$$

Replacing m by a subsequence m_n of n and denoting $K_n = K_{m_n,n}$ it follows that

$$\Pr[(\widehat{\xi}_{n,1}, ..., \widehat{\xi}_{n,K_n})' \in \Xi^{\text{Int}}_{K_n}] \to 1.$$

Lemma 6.1 now follows straightforwardly from the latter result.

8. Proof of Lemma 6.2

Recall that

$$\widehat{Z}_n(u) = \sum_{k=1}^{K_n} \left(\frac{1}{\sqrt{N}} \sum_{j=1}^{N} \bigtriangledown_k f_j(\boldsymbol{\xi}^0) \right) \eta_k(u)$$

where $n = n_N$ and

$$\eta_k(u) = 2^{-k}\sqrt{2}\cos(k\pi u). \tag{8.1}$$

Obviously, $\widehat{Z}_n(u)$ is an a.s. continuous random function on $[0, 1]$, or in other words, $\widehat{Z}_n(u)$ is a random element of the space $\mathcal{C}[0, 1]$ of continuous functions on $[0, 1]$ endowed with the sup metric $||f - g||_{\text{sup}} = \sup_{0\leq u\leq 1} |f(u) - g(u)|$.

Now let

$$\widetilde{Z}_N(u) = \sum_{k=1}^{\infty} \left(\frac{1}{\sqrt{N}} \sum_{j=1}^{N} \bigtriangledown_k f_j(\boldsymbol{\xi}^0) \right) \eta_k(u).$$

It follows from Assumptions 6.4 and 6.5 that

$$\begin{aligned} E\left[\sup_{0\leq u\leq 1} \left|\widetilde{Z}_N(u) - \widehat{Z}_n(u)\right|\right] &\leq \sqrt{2} \sum_{k=K_n+1}^{\infty} 2^{-k} E\left[\left|N^{-1/2} \sum_{j=1}^{N} \bigtriangledown_k f_j(\boldsymbol{\xi}^0)\right|\right] \\ &\leq \sqrt{2} \sum_{k=K_n+1}^{\infty} 2^{-k} \sqrt{E\left[\left(N^{-1/2} \sum_{j=1}^{N} \bigtriangledown_k f_j(\boldsymbol{\xi}^0)\right)^2\right]} \end{aligned}$$

$$
\begin{aligned}
&= \sqrt{2} \sum_{k=K_n+1}^{\infty} 2^{-k} \sqrt{E[(\triangledown_k f_1(\boldsymbol{\xi}^0))^2]} \\
&\to 0 \text{ as } n \to \infty,
\end{aligned}
$$

because by Schwarz inequality and Assumption 6.5,

$$
\begin{aligned}
\sum_{k=1}^{\infty} 2^{-k} \sqrt{E[(\triangledown_k f_1(\boldsymbol{\xi}^0))^2]} &\leq \sqrt{\sum_{k=1}^{\infty} 2^{-k}} \sqrt{\sum_{k=1}^{\infty} 2^{-k} E[(\triangledown_k f_1(\boldsymbol{\xi}^0))^2]} \\
&< \infty.
\end{aligned} \tag{8.2}
$$

Hence, $\widehat{Z}_n(u) = \widetilde{Z}_N(u) + o_p(1)$ uniformly in $u \in [0, 1]$. Therefore, it suffices to show that $\widetilde{Z}_N$ converges weakly to a zero mean Gaussian process Z on $[0, 1]$, denoted by $\widetilde{Z}_N \Rightarrow Z$.

Note that by (8.2),

$$
\begin{aligned}
\operatorname{var}(\widetilde{Z}_N(u)) &= E\left(\sum_{k=1}^{\infty} 2^{-k} \triangledown_k f_1(\boldsymbol{\xi}^0) \sqrt{2} \cos(k\pi u)\right)^2 \\
&\leq 2E\left[\left(\sum_{k=1}^{\infty} 2^{-k} |\triangledown_k f_1(\boldsymbol{\xi}^0)|\right)^2\right] \\
&= 2\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} 2^{-k} 2^{-m} E\left[|\triangledown_k f_1(\boldsymbol{\xi}^0)|.|\triangledown_m f_1(\boldsymbol{\xi}^0)|\right] \\
&\leq 2\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} 2^{-k} 2^{-m} \sqrt{E[\triangledown_k f_1(\boldsymbol{\xi}^0)^2]} \sqrt{E[\triangledown_m f_1(\boldsymbol{\xi}^0)^2]} \\
&= 2\left(\sum_{k=1}^{\infty} 2^{-k} \sqrt{E[\triangledown_k f_1(\boldsymbol{\xi}^0)^2]}\right)^2 < \infty.
\end{aligned}
$$

It is well-known (see for example Billingsley 1968) that $\widetilde{Z}_N \Rightarrow Z$ holds if and only if $\widetilde{Z}_N(u)$ is tight and the finite distributions of $\widetilde{Z}_N(u)$ converge to the corresponding finite distributions of $Z(u)$. As to the latter, we need to show that for arbitrary points $u_1, u_2, ..., u_M$ in $[0, 1]$,

$$
(\widetilde{Z}_N(u_1), ..., \widetilde{Z}_N(u_M))' \xrightarrow{\text{d}} (Z(u_1), ..., Z(u_M))'. \tag{8.3}
$$

This is easy to verify for the case $M = 2$:

$$\begin{aligned}
\begin{pmatrix} \widetilde{Z}_N(u_1) \\ \widetilde{Z}_N(u_2) \end{pmatrix} &= \sum_{k=1}^{\infty} \left(\frac{1}{\sqrt{N}} \sum_{j=1}^{N} \triangledown_k f_j(\boldsymbol{\xi}^0) \right) \begin{pmatrix} \eta_k(u_1) \\ \eta_k(u_2) \end{pmatrix} \\
&= \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \left(\sum_{k=1}^{\infty} \triangledown_k f_j(\boldsymbol{\xi}^0) \begin{pmatrix} \eta_k(u_1) \\ \eta_k(u_2) \end{pmatrix} \right) \\
&= \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \left(\sum_{k=1}^{\infty} 2^{-k} \triangledown_k f_j(\boldsymbol{\xi}^0) \begin{pmatrix} \sqrt{2}\cos(k\pi u_1) \\ \sqrt{2}\cos(k\pi u_2) \end{pmatrix} \right) \\
&\overset{\text{d}}{\to} \begin{pmatrix} Z(u_1) \\ Z(u_2) \end{pmatrix} \sim \mathcal{N}_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Gamma(u_1,u_1) & \Gamma(u_1,u_2) \\ \Gamma(u_1,u_2) & \Gamma(u_2,u_2) \end{pmatrix} \right]
\end{aligned}$$

where the latter follows from the standard central limit theorem, with $\Gamma(u_1, u_2)$ defined in Lemma 6.2.

Tightness is a generalization of the notion of stochastic boundedness to random functions, and since convergence in distribution implies stochastic boundedness, it follows similarly that tightness is a necessary condition for weak convergence. In particular, $\widetilde{Z}_N$ is tight if for an arbitrary $\varepsilon \in (0,1)$ there exists a compact set $K(\varepsilon) \in \mathcal{C}[0,1]$ such that $\inf_{N\geq 1} \Pr[\widetilde{Z}_N(.) \in K(\varepsilon)] > 1 - \varepsilon$.

As to the tightness of $\widetilde{Z}_N(u)$, it suffices to show that for arbitrary $\varepsilon > 0$,

$$\sup_{|u_1-u_2|\leq\varepsilon,\ u_1,u_2\in[0,1]} \left| \widetilde{Z}_N(u_1) - \widetilde{Z}_N(u_2) \right| = \varepsilon.O_p(1), \tag{8.4}$$

as then condition (8.3) of Theorem 8.2 in Billingsley (1968) holds.[2] This follows easily from the fact that by (8.1) and the mean value theorem,

$$\sup_{|u_1-u_2|\leq\varepsilon,\ u_1,u_2\in[0,1]} |\eta_k(u_1) - \eta_k(u_2)| \leq \varepsilon\sqrt{2}\pi k.2^{-k}$$

so that by Assumptions 6.4 and 6.5,

$$\begin{aligned}
& E\left[\sup_{|u_1-u_2|\leq\varepsilon,\ u_1,u_2\in[0,1]} \left| \widetilde{Z}_N(u_1) - \widetilde{Z}_N(u_2) \right| \right] \\
& \leq \sum_{k=1}^{\infty} E\left[\left| N^{-1/2} \sum_{j=1}^{N} \triangledown_k f_j(\boldsymbol{\xi}^0) \right| \right] \sup_{|u_1-u_2|\leq\varepsilon,\ u_1,u_2\in[0,1]} |\eta_k(u_1) - \eta_k(u_2)|
\end{aligned}$$

[2] Note that condition (8.2) of Theorem 8.2 in Billingsley (1968) follows from (8.3).

$$\leq \sum_{k=1}^{\infty} \sqrt{E[(\triangledown_k f_1(\boldsymbol{\xi}^0))^2]} \sup_{|u_1-u_2|\leq\varepsilon,\ u_1,u_2\in[0,1]} |\eta_k(u_1)-\eta_k(u_2)|$$

$$\leq \varepsilon\sqrt{2}\pi \sum_{k=1}^{\infty} k.2^{-k}\sqrt{E[(\triangledown_k f_1(\boldsymbol{\xi}^0))^2]}$$

$$\leq \varepsilon\sqrt{2}\pi \sqrt{\sum_{k=1}^{\infty} k.2^{-k}} \sqrt{\sum_{k=1}^{\infty} k.2^{-k} E[(\triangledown_k f_1(\boldsymbol{\xi}^0))^2]} = \varepsilon.O(1).$$

Finally, using the Cauchy-Schwarz and Liapounov inequalities and (8.1) it follows that

$$\begin{aligned}
\sup_{0\leq u_1\leq 1,\ 0\leq u_2\leq 1} |\Gamma(u_1,u_2)| &\leq 2\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} E[|(\nabla_k f_1(\boldsymbol{\xi}^0))(\nabla_m f_1(\boldsymbol{\xi}^0))|]2^{-m}2^{-k} \\
&\leq 2\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} \sqrt{E[(\nabla_k f_1(\boldsymbol{\xi}^0))^2]}\sqrt{E[(\nabla_m f_1(\boldsymbol{\xi}^0))^2]}2^{-m}2^{-k} \\
&= 2\left(\sum_{m=1}^{\infty} 2^{-m}\sqrt{E[(\nabla_m f_1(\boldsymbol{\xi}^0))^2]}\right)^2 \\
&= 2\left(\frac{\sum_{m=1}^{\infty} 2^{-m}\sqrt{E[(\nabla_m f_1(\boldsymbol{\xi}^0))^2]}}{\sum_{k=1}^{\infty} 2^{-k}}\right)^2 \left(\sum_{k=1}^{\infty} 2^{-k}\right)^2 \\
&\leq 2\frac{\sum_{m=1}^{\infty} 2^{-m} E[(\nabla_m f_1(\boldsymbol{\xi}^0))^2]}{\sum_{k=1}^{\infty} 2^{-k}} \left(\sum_{k=1}^{\infty} 2^{-k}\right)^2 \\
&= 2\left(\sum_{m=1}^{\infty} 2^{-m} E[(\nabla_m f_1(\boldsymbol{\xi}^0))^2]\right)\left(\sum_{k=1}^{\infty} 2^{-k}\right) < \infty
\end{aligned}$$

where the latter inequality follows from Assumption 6.5.

9. Proof of Lemma 6.4 (Corrections)

Equation (A.4) in the printed version of the paper is incorrect. It should be:

$$B'_{k,n}\Phi_k B_{k,n} = \begin{pmatrix} B'_{k,p}\Phi_k B_{k,p} & B'_{k,p}\Phi_k C_{k,n-p} \\ C'_{k,n-p}\Phi_k B_{k,p} & C'_{k,n-p}\Phi_k C_{k,n-p} \end{pmatrix}.$$

Similarly, the next equation after (A.4) should be

$$\lim_{k\to\infty} B'_{k,n}\Phi_k B_{k,n} = \begin{pmatrix} \lim_{k\to\infty} B'_{k,p}\Phi_k B_{k,p} & \lim_{k\to\infty} B'_{k,p}\Phi_k C_{k,n-p} \\ \lim_{k\to\infty} C'_{k,n-p}\Phi_k B_{k,p} & \lim_{k\to\infty} C'_{k,n-p}\Phi_k C_{k,n-p} \end{pmatrix}$$

10. Proof of Theorem 6.2: Continuation

To complete the proof of Theorem 6.2, it needs to be shown that

$$\lim_{n\to\infty} A_{n,n} = A, \tag{10.1}$$

where for $k \geq n$, $A_{k,n} = \int_0^1 a_{k,n}(u)a_{k,n}(u)'\mathrm{d}u$, with $a_{k,n}(u)$ defined in the proof of Lemma 6.4. Moreover, recall from Lemma 6.4 and its proof that

$$\lim_{n\to\infty}\lim_{k\to\infty}\int_0^1 a_{k,n}(u)a_{k,n}(u)'\mathrm{d}u = \lim_{n\to\infty}\int_0^1 a_n(u)a_n(u)'\mathrm{d}u = A$$

Also, recall from the proof of Lemma 6.4 that, with $B_{k,n}$ partitioned for $k > n$ as $B_{k,n} = (B_{k,p}, C_{k,n-p})$,

$$b^{(2)}_{k,n}(u) = -C'_{k,n-p}\varphi_k(u),\ b^{(2)}_{n,n}(u) = -C'_{n,n-p}\varphi_n(u) = -C'_{k,n-p}\begin{pmatrix}\varphi_n(u) \\ 0_{k-n}\end{pmatrix},$$

where $\varphi_k(u) = (\eta_1(u), ..., \eta_k(u))'$ with $\eta_k(u) = 2^{-k}\sqrt{2}\cos(k\pi u)$.

To prove (10.1), it will be shown first that

$$\lim_{n\to\infty}\lim_{k\to\infty}\int_0^1 \left(b^{(2)}_{k,n}(u) - b^{(2)}_{n,n}(u)\right)'\left(b^{(2)}_{k,n}(u) - b^{(2)}_{n,n}(u)\right)\mathrm{d}u = 0, \tag{10.2}$$

as follows. Partition $\varphi_k(u)$ as $\varphi_k(u) = (\varphi_n(u)', \varphi^*_{k-n}(u)')'$, so that

$$b^{(2)}_{n,n}(u) - b^{(2)}_{k,n}(u) = C'_{k,n-p}\begin{pmatrix}0_n \\ \varphi^*_{k-n}(u)\end{pmatrix}.$$

It is easy to verify that $\int_0^1 \varphi^*_{k-n}(u)\varphi^*_{k-n}(u)'\mathrm{d}u = 2^{-2n}\Phi_{k-n}$, where

$$\Phi_{k-n} = \int_0^1 \varphi_{k-n}(u)\varphi_{k-n}(u)'\mathrm{d}u,$$

hence

$$\begin{aligned}\mu_{k,n} &= \int_0^1 \left(b^{(2)}_{k,n}(u) - b^{(2)}_{n,n}(u)\right)' \left(b^{(2)}_{k,n}(u) - b^{(2)}_{n,n}(u)\right) \mathrm{d}u \\ &= 2^{-2n} C'_{k,n-p} \begin{pmatrix} O_{n,n} & O_{n,k-n} \\ O_{k-n,n} & \Phi_{k-n} \end{pmatrix} C_{k,n-p} \\ &= 2^{-2n} C^{*\prime}_{k-n,n-p} \Phi_{k-n} C^{*}_{k-n,n-p}\end{aligned}$$

where $C^*_{k-n,n-p}$ is the matrix of the last $k-n$ rows of $C_{k,n-p}$, i.e.,

$$C^*_{k,n-p} = \begin{pmatrix} E[\nabla_{n+1,p+1} f_1(\boldsymbol{\xi}^0)] & \cdots & E[\nabla_{n+1,n} f_1(\boldsymbol{\xi}^0)] \\ \vdots & \ddots & \vdots \\ E[\nabla_{k,p+1} f_1(\boldsymbol{\xi}^0)] & \cdots & E[\nabla_{k,n} f_1(\boldsymbol{\xi}^0)] \end{pmatrix}.$$

Then for some constant $c > 0$ and a sufficiently large n,

$$\begin{aligned}\mu_{k,n} &= 2^{-2n} \sum_{m=1}^{k-n} 2^{-2m} \sum_{s=p+1}^{n} \left(E[\nabla_{n+m,s} f_1(\boldsymbol{\xi}^0)]\right)^2 \\ &\leq 2^{-2n} n^c \sum_{m=1}^{k-n} 2^{-2m}(n+m)^c \sum_{s=1}^{n} \left(E[\nabla_{n+m,s} f_1(\boldsymbol{\xi}^0)]\right)^2 (n+m)^{-c} s^{-c} \\ &\leq 2^{-2} 2^{-2n} n^c (n+1)^c \sum_{m=1}^{k-n} \sum_{s=1}^{n} \left(E[\nabla_{n+m,s} f_1(\boldsymbol{\xi}^0)]\right)^2 (n+m)^{-c} s^{-c} \\ &\leq 2^{-2} 2^{-2n} n^c (n+1)^c \sum_{m=1}^{k} \sum_{s=1}^{n} \left(E[\nabla_{m,s} f_1(\boldsymbol{\xi}^0)]\right)^2 m^{-c} s^{-c} \\ &\leq 2^{-2} 2^{-2n} n^c (n+1)^c \sum_{m=1}^{k} \sum_{s=1}^{n} \left(\left|E[\nabla_{m,s} f_1(\boldsymbol{\xi}^0)]\right| m^{-c/2} s^{-c/2}\right)^2 \\ &\leq 2^{-2} 2^{-2n} n^c (n+1)^c \left(\sum_{m=1}^{k} \sum_{s=1}^{n} m^{-c/2} s^{-c/2} \left|E[\nabla_{m,s} f_1(\boldsymbol{\xi}^0)]\right|\right)^2\end{aligned}$$

where the second inequality is due to the easy inequality

$$\max_{m \geq 1} 2^{-2m}(n+m)^c \leq \max_{x \geq 1} 2^{-2x}(n+x)^c = 2^{-2}(n+1)^c$$

for $n \geq c.(2\ln(2))^{-1} - 1$.

By part (a) of Assumption 6.6, $\sum_{k=1}^{\infty}\sum_{m=1}^{\infty}(k.m)^{-2-\tau}E[|\nabla_{k,m}f_1(\boldsymbol{\xi}^0)|] < \infty$, hence for $c = 4 + 2\tau$,

$$\begin{aligned}\mu_{k,n} &\leq 2^{-2}2^{-2n}n^{4+2\tau}(n+1)^{4+2\tau}\left(\sum_{j=1}^{\infty}\sum_{m=1}^{\infty}(j.m)^{-2-\tau}E[|\nabla_{j,m}f_1(\boldsymbol{\xi}^0)|]\right)^2 \\ &= o\left(2^{-2n}(n+1)^{8+4\tau}\right) = o(1)\end{aligned}$$

Thus, $\lim_{n\to\infty}\sup_{k\geq n}\mu_{k,n} = 0$, which proves (10.2).

By a similar argument it can be shown that

$$\lim_{k\to\infty}\int_0^1\left(b_k^{(1)}(u) - b(u)\right)'\left(b_k^{(1)}(u) - b(u)\right)\mathrm{d}u = 0, \tag{10.3}$$

where $b_k^{(1)}(u) = -B'_{k,p}\varphi_k(u)$ and $b(u)$ is defined in Lemma 6.3. It follows now from (10.2), (10.3) and Theorem B.1 that

$$\lim_{n\to\infty}\int_0^1(a_{n,n}(u) - a_n(u))'(a_{n,n}(u) - a_n(u))\mathrm{d}u = 0,$$

which implies (10.1). This completes the proof of Theorem 6.2.

11. Proof of Lemma 6.5

First, note that

$$\begin{aligned}&\left|\frac{1}{N}\sum_{j=1}^{N}\left((\triangledown_k f_j(\widehat{\boldsymbol{\xi}}_n))(\triangledown_m f_j(\widehat{\boldsymbol{\xi}}_n)) - (\triangledown_k f_j(\boldsymbol{\xi}^0))(\triangledown_m f_j(\boldsymbol{\xi}^0))\right)\right| \\ &\quad\leq \frac{1}{N}\sum_{j=1}^{N}|\triangledown_k f_j(\widehat{\boldsymbol{\xi}}_n) - \triangledown_k f_j(\boldsymbol{\xi}^0)|.|\triangledown_m f_j(\widehat{\boldsymbol{\xi}}_n) - \triangledown_m f_j(\boldsymbol{\xi}^0)| \\ &\quad+ \frac{1}{N}\sum_{j=1}^{N}|\triangledown_k f_j(\boldsymbol{\xi}^0)|.|\triangledown_m f_j(\widehat{\boldsymbol{\xi}}_n) - \triangledown_m f_j(\boldsymbol{\xi}^0)| \\ &\quad+ \frac{1}{N}\sum_{j=1}^{N}|\triangledown_m f_j(\boldsymbol{\xi}^0)|.|\triangledown_k f_j(\widehat{\boldsymbol{\xi}}_n) - \triangledown_k f_j(\boldsymbol{\xi}^0)|\end{aligned}$$

$$
\begin{aligned}
&\leq \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_k f_j(\widehat{\boldsymbol{\xi}}_n) - \nabla_k f_j(\boldsymbol{\xi}^0))^2} \\
&\times \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_m f_j(\widehat{\boldsymbol{\xi}}_n) - \nabla_m f_j(\boldsymbol{\xi}^0))^2} \\
&+ \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_k f_j(\boldsymbol{\xi}^0))^2}\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_m f_j(\widehat{\boldsymbol{\xi}}_n) - \nabla_m f_j(\boldsymbol{\xi}^0))^2} \\
&+ \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_m f_j(\boldsymbol{\xi}^0))^2}\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_k f_j(\widehat{\boldsymbol{\xi}}_n) - \nabla_k f_j(\boldsymbol{\xi}^0))^2}
\end{aligned}
$$

Then with $\eta_k(u) = 2^{-k}\sqrt{2}\cos(k\pi u)$, and

$$
\widehat{d}_N = \frac{1}{N}\sum_{j=1}^{N}\left(\sum_{k=1}^{\infty} 2^{-k}(\nabla_k f_j(\widehat{\boldsymbol{\xi}}_n) - \nabla_k f_j(\boldsymbol{\xi}^0))^2\right)
$$

we have

$$
\begin{aligned}
&\sup_{u_1,u_2\in[0,1]}\left|\widehat{\Gamma}_n(u_1,u_2) - \Gamma_n(u_1,u_2)\right| \\
&\leq 2\left(\sum_{k=1}^{\infty} 2^{-k}\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_k f_j(\widehat{\boldsymbol{\xi}}_n) - \nabla_k f_j(\boldsymbol{\xi}^0))^2}\right)^2 \\
&+ 4\left(\sum_{k=1}^{\infty} 2^{-k}\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_k f_j(\widehat{\boldsymbol{\xi}}_n) - \nabla_k f_j(\boldsymbol{\xi}^0))^2}\right) \\
&\qquad \times \left(\sum_{k=1}^{\infty} 2^{-k}\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nabla_k f_j(\boldsymbol{\xi}^0))^2}\right) \\
&\leq \left(2\widehat{d}_N + 4\sqrt{\widehat{d}_N}\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\sum_{k=1}^{\infty} 2^{-k}(\nabla_k f_j(\boldsymbol{\xi}^0))^2\right)}\right)\sum_{k=1}^{\infty} 2^{-k}
\end{aligned}
$$

where the latter inequality follows from Schwarz inequality. Note that by Assumption 6.5,

$$\sum_{k=1}^{n} 2^{-k} \frac{1}{N} \sum_{j=1}^{N} (\triangledown_k f_j(\boldsymbol{\xi}^0))^2 = O_p(1).$$

Therefore, it suffices to show that $\widehat{d}_N = o_p(1)$, as follows. For arbitrary $\varepsilon > 0$,

$$\Pr\left[\widehat{d}_N = \widehat{d}_N . I\left(||\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}^0||_\ell \le \varepsilon\right)\right] = \Pr\left[||\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}^0||_\ell \le \varepsilon\right] \to 1$$

and

$$E\left[\widehat{d}_N . I\left(||\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}^0||_\ell \le \varepsilon\right)\right] \le \sum_{k=1}^{\infty} 2^{-k} E\left[\sup_{||\boldsymbol{\xi} - \boldsymbol{\xi}^0||_\ell \le \varepsilon} (\triangledown_k f_1(\boldsymbol{\xi}) - \triangledown_k f_1(\boldsymbol{\xi}^0))^2\right]$$

By Assumption 6.9 the latter can be made arbitrarily small. It follows now straightforwardly that $\widehat{d}_N = o_p(1)$.

12. Proof of Theorem 6.5

Recall that for fixed $s \in \mathbb{N}$,

$$\sqrt{N}\left(\widehat{\delta}_{n_N,1} - \delta_{0,1}, ..., \widehat{\delta}_{n_N,s} - \delta_{0,s}\right)' \xrightarrow{\mathrm{d}} \mathcal{N}_s(0, \Lambda_s) \tag{12.1}$$

for some positive-definite $s \times s$ matrix Λ_s.

We can write

$$\begin{aligned}
&\sqrt{N}\left(h(u|\pi_s \widehat{\boldsymbol{\delta}}_{n_N}) - h(u|\pi_s \boldsymbol{\delta}^0)\right) \\
&= \frac{\sum_{m=1}^{s} \sqrt{N}\left(\widehat{\delta}_{n_N,m} - \delta_{0,m}\right)\sqrt{2}\cos(m\pi u)}{1 + \sum_{m=1}^{s} \widehat{\delta}_{n_N,m}^2} \\
&\quad \times \left(2 + \sum_{m=1}^{s}\left(\widehat{\delta}_{n_N,m} + \delta_{0,m}\right)\sqrt{2}\cos(m\pi u)\right) \\
&\quad - \frac{\sum_{m=1}^{s} \sqrt{N}\left(\widehat{\delta}_{n_N,m} - \delta_{0,m}\right)\left(\widehat{\delta}_{n_N,m} + \delta_{0,m}\right)}{1 + \sum_{m=1}^{s} \widehat{\delta}_{n_N,m}^2} h(u|\pi_s \boldsymbol{\delta}^0)
\end{aligned}$$

$$= 2\frac{\sum_{m=1}^{s}\sqrt{N}\left(\widehat{\delta}_{n_N,m}-\delta_{0,m}\right)\sqrt{2}\cos(m\pi u)}{1+\sum_{m=1}^{s}\delta_{0,m}^2}$$
$$\times\left(1+\sum_{m=1}^{s}\delta_{0,m}\sqrt{2}\cos(m\pi u)\right)$$
$$-2\frac{\sum_{m=1}^{s}\sqrt{N}\left(\widehat{\delta}_{n_N,m}-\delta_{0,m}\right)\delta_{0,m}}{1+\sum_{m=1}^{s}\delta_{0,m}^2}h(u|\pi_s\boldsymbol{\delta}^0)+o_p(1),$$

where the $o_p(1)$ is uniform in $u\in[0,1]$. The latter equality follows from the fact that by (12.1), $\sum_{m=1}^{s}\sqrt{N}|\widehat{\delta}_{n_N,m}-\delta_{0,m}|=O_p(1)$ and $\text{plim}_{N\to\infty}\widehat{\delta}_{n_N,m}=\delta_{0,m}$ for $m=1,2,...,s$.

Denote $\widehat{Z}_s=\sqrt{N}\left(\widehat{\delta}_1-\delta_{0,1},...,\widehat{\delta}_q-\delta_{0,s}\right)$. Using the notations $\kappa_s(u)$, δ_s and $\omega_s(u)$ in Theorem 6.5 we can write

$$\widehat{\Phi}_s(u)=\sqrt{N}\left(h(u|\pi_s\widehat{\boldsymbol{\delta}}_{n_N})-h(u|\pi_s\boldsymbol{\delta}^0)\right)=\frac{2}{1+\delta_s{}'\delta_s}\widehat{Z}_s'\omega_s(u)+o_p(1)$$

It is now easy to verify from (12.1) that $\widehat{\Phi}_s(u)$ is tight and converges weakly to a zero-mean Gaussian process $\widetilde{\Phi}_s(u)$, similar to Lemma 6.2, with covariance function

$$E[\widetilde{\Phi}_s(u_1)\widetilde{\Phi}_s(u_2)]=\frac{4}{(1+\delta_s'\delta_s)^2}\omega_s(u_1)'\Lambda_s\omega_s(u_1).$$

13. Proof of Theorem 6.6

Denote

$$\widehat{\phi}_N(u) = 1+\sum_{m=1}^{n_N}\widehat{\delta}_{n_N,m}\sqrt{2}\cos(m\pi u),$$
$$\phi_0(u) = 1+\sum_{m=1}^{\infty}\delta_{0,m}\sqrt{2}\cos(m\pi u),$$

and recall that

$$h(u|\widehat{\boldsymbol{\delta}}_{n_N}) = \widehat{\phi}_N(u)^2/\int_0^1\widehat{\phi}_N(v)^2\mathrm{d}v,$$
$$h_0(u) = h(u|\boldsymbol{\delta}^0)=\phi_0(u)^2/\int_0^1\phi_0(v)^2\mathrm{d}v$$

Next, let

$$\begin{aligned}
\widehat{T}_N(u) &= \sqrt{N}\left(\widehat{\phi}_N(u) - \phi_0(u)\right) \\
&= \sum_{m=1}^{n_N} \sqrt{N}\left(\widehat{\delta}_{n_N,m} - \delta_{0,m}\right)\sqrt{2}\cos(m\pi u) \\
&\quad -\sqrt{N}\sum_{m=n_N+1}^{\infty} \delta_{0,m}\sqrt{2}\cos(m\pi u) \\
&= \sum_{m=1}^{n_N} \sqrt{N}\left(\widehat{\delta}_{n_N,m} - \delta_{0,m}\right)\sqrt{2}\cos(m\pi u) + o(1) \qquad (13.1)
\end{aligned}$$

where the $o(1)$ term is uniform in $u \in [0,1]$. The latter follows from

$$\begin{aligned}
\left|\sqrt{N}\sum_{m=n_N+1}^{\infty} \delta_{0,m}\sqrt{2}\cos(m\pi u)\right| &\leq \sqrt{2N}\sum_{m=n_N+1}^{\infty} |\delta_{0,m}| \\
&\leq \frac{\sqrt{2N}}{n_N^{\ell}}\sum_{m=n_N+1}^{\infty} m^{\ell}|\delta_{0,m}| = o(1) \qquad (13.2)
\end{aligned}$$

for

$$n_N \propto N^{1/(2\ell)}. \qquad (13.3)$$

See Remark B.4 following Assumption 6.3 for (13.3).

First note that the functions $\widehat{T}_N(u)$ are random elements of the Hilbert space $L_0^2(0,1) = \mathrm{span}(\{\sqrt{2}\cos(m\pi u)\}_{m=1}^{\infty})$.

According to Van der Vaart and Wellner (1996, Ch. 1.8) [VW hereafter], $\widehat{T}_N(u)$ is asymptotically finite-dimensional if for each $\eta, \varepsilon > 0$ there exists a finite subset $\mathbb{I} \subset \mathbb{N}$ such that

$$\begin{aligned}
&\limsup_{N\to\infty} \Pr\left[\sum_{i\in\mathbb{I}}\left(\int_0^1 \widehat{T}_N(u)\sqrt{2}\cos(i.\pi u)\,\mathrm{d}u\right)^2 > \eta\right] \\
&\qquad = \limsup_{N\to\infty} \Pr\left[N\sum_{i\in\mathbb{I}}\left(\widehat{\delta}_{n_N,i} - \delta_{0,i}\right)^2 > \eta\right] < \varepsilon. \qquad (13.4)
\end{aligned}$$

Since by (12.1) $N\sum_{i\in\mathbb{I}}\left(\widehat{\delta}_{n_N,i} - \delta_{0,i}\right)^2$ converges in distribution to a quadratic form of zero-mean jointly normal random variables, (13.4) holds, hence $\widehat{T}_N(u)$ is asymptotically finite-dimensional. Then it follows from Lemma 1.8.1 in VW that $\widehat{T}_N(u)$ is asymptotically tight if for each $m \in \mathbb{N}$,

$$\int_0^1 \widehat{T}_N(u)\sqrt{2}\cos(m\pi u)\,\mathrm{d}u = N\left(\widehat{\delta}_{n_N,m} - \delta_{0,m}\right)^2 I(m \leq n_N) - \delta_{0,m}^2 I(m > n_N)$$

is asymptotically tight, which also holds by (12.1). Moreover, by Lemma 1.8.2 in VW, $\widehat{T}_N(u)$ is then asymptotically measurable. It follows now from Prohorov's theorem [van der Vaart (1998, Th.18.12(ii))] that there exists a subsequence N_j and a tight random element $\widetilde{T} \in L_0^2(0,1)$ such that

$$\widehat{T}_{N_j}(u) = \sqrt{N_j}\left(\widehat{\phi}_{N_j}(u) - \phi_0(u)\right) \rightsquigarrow \widetilde{T}(u) \text{ as } j \to \infty. \tag{13.5}$$

Moreover, since $\widehat{\phi}_{N_j}(u) \xrightarrow{\mathrm{p}} \phi_0(u)$ uniformly on $[0,1]$, it follows that

$$\operatorname*{plim}_{j\to\infty}\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v} = \sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}.$$

Next, consider the following sequence of random variables:

$$\begin{aligned}
&\sqrt{N_j}\left(\frac{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}} - 1\right) \\
&= \sqrt{N_j}\frac{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} - \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}} \\
&\qquad \times \frac{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} + \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}}{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} + \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}} \\
&= -\frac{\int_0^1 \sqrt{N_j}\left(\widehat{\phi}_{N_j}(v) - \phi_0(v)\right)\left(\widehat{\phi}_{N_j}(v) + \phi_0(v)\right)\mathrm{d}v}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\left(\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} + \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\right)} \\
&= \frac{-1}{\sqrt{N_j}}\frac{\int_0^1 \left(\sqrt{N_j}\left(\widehat{\phi}_{N_j}(v) - \phi_0(v)\right)\right)^2\mathrm{d}v}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\left(\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} + \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\right)}
\end{aligned}$$

$$- \frac{2\int_0^1 \sqrt{N_j}\left(\widehat{\phi}_{N_j}(v) - \phi_0(v)\right)\phi_0(v)\mathrm{d}v}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\left(\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} + \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\right)}$$

Since by the continuous mapping theorem,

$$\frac{\int_0^1 \left(\sqrt{N_j}\left(\widehat{\phi}_{N_j}(v) - \phi_0(v)\right)\right)^2 \mathrm{d}v}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\left(\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} + \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\right)} \xrightarrow{\mathrm{d}} \frac{\int_0^1 \widetilde{T}(v)^2\mathrm{d}v}{2\int_0^1 \phi_0(v)^2\mathrm{d}v}$$

and

$$\frac{2\int_0^1 \sqrt{N_j}\left(\widehat{\phi}_{N_j}(v) - \phi_0(v)\right)\phi_0(v)\mathrm{d}v}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\left(\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v} + \sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}\right)} \xrightarrow{\mathrm{d}} \frac{\int_0^1 \widetilde{T}(v)\phi_0(v)\mathrm{d}v}{\int_0^1 \phi_0(v)^2\mathrm{d}v}$$

it follows that

$$\sqrt{N_j}\left(\frac{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}} - 1\right) \xrightarrow{\mathrm{d}} -\frac{\int_0^1 \widetilde{T}(v)\phi_0(v)\mathrm{d}v}{\int_0^1 \phi_0(v)^2\mathrm{d}v}$$

Hence,

$$\begin{aligned}
&\sqrt{N_j}\left(\frac{\widehat{\phi}_{N_j}(u)}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}} - \frac{\phi_0(u)}{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}}\right) = \frac{\sqrt{N_j}(\widehat{\phi}_{N_j}(u) - \phi_0(u))}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}} \\
&+\sqrt{N_j}\left(\frac{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2\mathrm{d}v}} - 1\right)\frac{\phi_0(u)}{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}} \\
&\rightsquigarrow \frac{\widetilde{T}(u)}{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}} - \frac{\int_0^1 \widetilde{T}(v)\phi_0(v)\mathrm{d}v}{\int_0^1 \phi_0(v)^2\mathrm{d}v} \times \frac{\phi_0(u)}{\sqrt{\int_0^1 \phi_0(v)^2\mathrm{d}v}} \\
&= \left(\int_0^1 \sqrt{h_0(v)}\mathrm{d}v\right) \times \left(\widetilde{T}(u) - \sqrt{h_0(u)}\int_0^1 \widetilde{T}(v)\sqrt{h_0(v)}\mathrm{d}v\right)
\end{aligned}$$

The latter follows from the condition that $h_0(u)$ is uniformly continuous and positive on $[0, 1]$, so that

$$\phi_0(u) > 0 \text{ on } [0, 1]. \tag{13.6}$$

Hence,

$$\sqrt{h_0(u)} = \frac{\phi_0(u)}{\sqrt{\int_0^1 \phi_0(v)^2 \mathrm{d}v}}, \int_0^1 \sqrt{h_0(v)}\mathrm{d}v = \frac{1}{\sqrt{\int_0^1 \phi_0(v)^2 \mathrm{d}v}}.$$

Moreover, because $\widehat{\phi}_{N_j}(u) \xrightarrow{\mathrm{p}} \phi_0(u)$ uniformly on $[0,1]$, it follows from (13.6) that

$$\begin{aligned} \lim_{j\to\infty} \Pr\left[\sup_{0<u<1}\left|\sqrt{h(u|\widehat{\boldsymbol{\delta}}_{n_{N_j}})} - \frac{\widehat{\phi}_{N_j}(u)}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2 \mathrm{d}v}}\right| = 0\right] &= 1, \\ \lim_{j\to\infty} \Pr\left[\int_0^1 \sqrt{h(v|\widehat{\boldsymbol{\delta}}_{n_{N_j}})}\mathrm{d}v = \frac{1}{\sqrt{\int_0^1 \widehat{\phi}_{N_j}(v)^2 \mathrm{d}v}}\right] &= 1. \end{aligned}$$

Thus

$$\begin{aligned} &\sqrt{N_j}\left(\sqrt{h(u|\widehat{\boldsymbol{\delta}}_{n_{N_j}})} - \sqrt{h_0(u)}\right) \\ &\rightsquigarrow \left(\int_0^1 \sqrt{h_0(v)}\mathrm{d}v\right)\left(\widetilde{T}(u) - \sqrt{h_0(u)}\int_0^1 \widetilde{T}(v)\sqrt{h_0(v)}\mathrm{d}v\right) \\ &= \widetilde{\Omega}(u), \end{aligned} \tag{13.7}$$

say.

To prove that $\widetilde{T}(u)$ is zero-mean Gaussian, let

$$f(u) = \sum_{m=1}^{\infty} \gamma_m \sqrt{2}\cos(m\pi u), \sum_{m=1}^{\infty} \gamma_m^2 < \infty, \tag{13.8}$$

be an arbitrary function in $L_0^2(0,1)$, and let for $s \in \mathbb{N}$,

$$f_s(u) = \sum_{m=1}^{s} \gamma_m \sqrt{2}\cos(m\pi u)$$

It follows from (12.1), (13.1), (13.5) and the continuous mapping theorem that, with $n_{N_j} \geq s$ and $\omega_s = (\gamma_1, ..., \gamma_s)'$,

$$\begin{aligned} \int_0^1 \widehat{T}_{N_j}(u) f_s(u)\mathrm{d}u &= \sum_{m=1}^{s} \sqrt{N_j}\left(\widehat{\delta}_{n_{Nj},m} - \delta_{0,m}\right)\gamma_m \\ &\xrightarrow{\mathrm{d}} \int_0^1 \widetilde{T}(u) f_s(u)\mathrm{d}u \sim \mathcal{N}(0, \omega_s'\Lambda_s\omega_s) \end{aligned} \tag{13.9}$$

Moreover,

$$
\begin{aligned}
&\left|\int_0^1 \widetilde{T}(u)f(u)\mathrm{d}u - \int_0^1 \widetilde{T}(u)f_s(u)\mathrm{d}u\right| \\
&\quad \le \sqrt{\int_0^1 \widetilde{T}(u)^2\mathrm{d}u}\sqrt{\int_0^1 \left(f(u)-f_s(u)\right)^2\mathrm{d}u} \\
&\quad = \sqrt{\int_0^1 \widetilde{T}(u)^2\mathrm{d}u}\sqrt{\sum_{m=s+1}^{\infty}\gamma_m^2} \xrightarrow{\mathrm{p}} 0 \text{ as } s\to\infty
\end{aligned}
$$

hence

$$\int_0^1 \widetilde{T}(u)f(u)\mathrm{d}u \sim \mathcal{N}\left(0, \lim_{s\to\infty}\omega_s'\Lambda_s\omega_s\right)$$

Since $\widetilde{T} \in L_0^2(0,1)$ with probability 1 it can be written as

$$\widetilde{T}(u) = \sum_{m=1}^{\infty}\varepsilon_m\sqrt{2}\cos(m\pi u)\,,\ \sum_{m=1}^{\infty}\varepsilon_m^2 < \infty \text{ a.s.},$$

so that

$$
\begin{aligned}
\int_0^1 \widetilde{T}(u)f_s(u)\mathrm{d}u &= \sum_{m=1}^{s}\varepsilon_m\gamma_m \sim \mathcal{N}(0,\omega_s'\Lambda_s\omega_s) \\
\int_0^1 \widetilde{T}(u)f(u)\mathrm{d}u &= \sum_{m=1}^{\infty}\varepsilon_m\gamma_m \sim \mathcal{N}(0,\lim_{s\to\infty}\omega_s'\Lambda_s\omega_s)
\end{aligned}
$$

Therefore, by Lemma 1.8.3 in VW, the ε_m's are zero-mean jointly normal with $\mathrm{Var}((\varepsilon_1,...,\varepsilon_s)') = \Lambda_s$ for all $s \in \mathbb{N}$, so that $\widetilde{T}(u)$ is zero-mean Gaussian process, and so is the process $\widetilde{\Omega}(u)$ in (13.7).

Finally, note that by Theorem 1.8.4 in VW we can only replace N_j by N itself if for all $f \in L_0^2(0,1)$, $\int_0^1 \widehat{T}_N(u)f(u)\mathrm{d}u \xrightarrow{\mathrm{d}} \int_0^1 \widetilde{T}(u)f(u)\mathrm{d}u$, which by (13.1) and (13.8) is equivalent to

$$\sum_{m=1}^{n_N}\sqrt{N}(\widehat{\delta}_{n_N,m}-\delta_{0,m})\gamma_m \xrightarrow{\mathrm{d}} \sum_{m=1}^{\infty}\varepsilon_m\gamma_m$$

for all sequences $\{\gamma_m\}_{m=1}^{\infty}$ satisfying $\sum_{m=1}^{\infty}\gamma_m^2 < \infty$.

14. Proof of Lemma 7.1

For the SNP Logit model in log-likelihood form without penalty function the function $f(Z, \boldsymbol{\xi})$ takes the form

$$\begin{aligned} f(Z, \boldsymbol{\xi}) &= f(Z, (\theta, \boldsymbol{\delta})) \\ &= Y \ln\left(H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) + (1-Y)\ln\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right), \end{aligned}$$

where $Z = (Y, X')'$ and $\boldsymbol{\xi} = (\theta, \boldsymbol{\delta})$. Denote

$$\widetilde{X} = (1, X')' = (\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_p)'.$$

With ∇_k indicating the derivative to the k-th parameter, the first derivatives of $f(Z, \boldsymbol{\xi})$ are

$$\begin{aligned} \nabla_k f(Z, \boldsymbol{\xi}) &= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) \\ &\quad \times \frac{G'(\widetilde{X}'\theta)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)} \widetilde{X}_k \\ &= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) \\ &\quad \times \frac{G(\widetilde{X}'\theta)(1 - G(\widetilde{X}'\theta))}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)} \widetilde{X}_k \\ &= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\widetilde{X}_k \quad (14.1) \end{aligned}$$

for $k = 1, ..., p$, where

$$\phi(u|\boldsymbol{\delta}) = \frac{u(1-u)}{H(u|\boldsymbol{\delta})(1 - H(u|\boldsymbol{\delta}))}, \tag{14.2}$$

and

$$\begin{aligned} \nabla_{p+m} f(Z, \boldsymbol{\xi}) &= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) \\ &\quad \times \frac{\nabla_m H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)} \\ &= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) \psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) \quad (14.3) \end{aligned}$$

for $m \in \mathbb{N}$, where

$$\psi_m(u|\boldsymbol{\delta}) = \frac{\bigtriangledown_m H(u|\boldsymbol{\delta})}{u(1-u)}. \tag{14.4}$$

14.1. Part (a) for $k \leq p$

It follows from (14.1) that for $k = 1, 2, ..., p$,

$$\begin{aligned}
&\bigtriangledown_k f(Z, \boldsymbol{\xi}) - \bigtriangledown_k f(Z, \boldsymbol{\xi}^0) \\
&= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) \\
&\times \Big\{h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(\phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) - \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0)\right)\widetilde{X}_k \\
&+h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(\phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0) - \phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)\right)\widetilde{X}_k \\
&+\left(h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) - h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0\right)\right)\phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)\widetilde{X}_k \\
&+\left(h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0\right) - h\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)\right)\phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)\widetilde{X}_k\Big\} \\
&+\left(H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0\right) - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)h\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)\phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)\widetilde{X}_k \\
&+\left(H\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right) - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0\right)\right)h\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)\phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)\widetilde{X}_k \\
&\times h\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)\phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)\widetilde{X}_k
\end{aligned} \tag{14.5}$$

so that with

$$U_X = G(\widetilde{X}'\theta_0) \tag{14.6}$$

and $\boldsymbol{\delta}_n^0 = \pi_n \boldsymbol{\delta}^0$,

$$\begin{aligned}
&\left|\bigtriangledown_k f(Z, \boldsymbol{\xi}_{n+p}^0) - \bigtriangledown_k f(Z, \boldsymbol{\xi}^0)\right| \\
&\quad \leq \sup_{\boldsymbol{\delta} \in \Delta_0(M)} \sup_{0 \leq u \leq 1} h\left(u|\boldsymbol{\delta}\right)\left|\phi(U_X|\boldsymbol{\delta}_n^0) - \phi(U_X|\boldsymbol{\delta}^0)\right|.\max(1, ||X||) \\
&\quad \left(1 + 2\sup_{0 \leq u \leq 1} h(u|\boldsymbol{\delta}^0)\right)\phi(U_X|\boldsymbol{\delta}^0)\sup_{0 \leq u \leq 1}|h(u|\boldsymbol{\delta}_n^0) - h(u|\boldsymbol{\delta}^0)|.\max(1, ||X||)
\end{aligned}$$

By Lemma 5.1 there exists a constant $C > 0$ such that

$$\sup_{0 \leq u \leq 1}|h\left(u|\boldsymbol{\delta}\right) - h(u|\boldsymbol{\delta}^0)| \leq C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0. \tag{14.7}$$

Moreover, note that for some constant C,

$$
\begin{aligned}
&\left|\phi(u|\boldsymbol{\delta}) - \phi(u|\boldsymbol{\delta}^0)\right| \\
&= u\,(1-u)\,\frac{\left|H(u|\boldsymbol{\delta}^0)\left(1 - H(u|\boldsymbol{\delta}^0)\right) - H(u|\boldsymbol{\delta})\,(1 - H(u|\boldsymbol{\delta}))\right|}{H(u|\boldsymbol{\delta})\,(1 - H(u|\boldsymbol{\delta}))\,H(u|\boldsymbol{\delta}^0)\left(1 - H(u|\boldsymbol{\delta}^0)\right)} \\
&\leq 2\left(\sup_{\boldsymbol{\delta}\in\Delta_0(M)}\sup_{0\leq v\leq 1} h\,(v|\boldsymbol{\delta})\right).\phi(u|\boldsymbol{\delta}).\phi(u|\boldsymbol{\delta}^0)\sup_{0\leq v\leq 1}|h(v|\boldsymbol{\delta}^0) - h(v|\boldsymbol{\delta})| \\
&\leq 2.C.\phi(u|\boldsymbol{\delta}).\phi(u|\boldsymbol{\delta}^0).||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0
\end{aligned}
$$

where the second inequality is due (14.7).

At this point I will now use Assumption 2.3, which together with part (d) of Assumption 2.1 implies that

$$
\sup_{0\leq u\leq 1}\phi(u|\boldsymbol{\delta}^0) \leq \left(\inf_{0\leq u\leq 1} h(u|\boldsymbol{\delta}^0)\right)^{-2} < \infty. \tag{14.8}
$$

Moreover, by Lemma 5.1 there exists a $d > 0$ such that $\sup_{0\leq u\leq 1}|h(u|\boldsymbol{\delta}) - h(u|\boldsymbol{\delta}^0)| < 0.5.\inf_{0\leq u\leq 1} h(u|\boldsymbol{\delta}^0)$ if $||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 < d$, so that

$$
\begin{aligned}
\sup_{0\leq u\leq 1}\phi(u|\boldsymbol{\delta}) &\leq \left(\inf_{0\leq u\leq 1} h(u|\boldsymbol{\delta})\right)^{-2} \\
&\leq \left(\inf_{0\leq u\leq 1}\left(h(u|\boldsymbol{\delta}) - h(u|\boldsymbol{\delta}^0)\right) + \inf_{0\leq u\leq 1} h(u|\boldsymbol{\delta}^0)\right)^{-2} \\
&\leq \left(\inf_{0\leq u\leq 1} h(u|\boldsymbol{\delta}^0) - \sup_{0\leq u\leq 1}|h(u|\boldsymbol{\delta}) - h(u|\boldsymbol{\delta}^0)|\right)^{-2} \\
&\leq 4\left(\inf_{0\leq u\leq 1} h(u|\boldsymbol{\delta}^0)\right)^{-2} < \infty.
\end{aligned} \tag{14.9}
$$

Hence, there exist constants $C > 0$ and $d > 0$ such that

$$
\sup_{0\leq u\leq 1}\left|\phi(u|\boldsymbol{\delta}) - \phi(u|\boldsymbol{\delta}^0)\right| \leq C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 \text{ if } ||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 < d. \tag{14.10}
$$

It follows now straightforwardly that

$$
\begin{aligned}
&\max_{1\leq k\leq p}\left|\triangledown_k f(Z,\boldsymbol{\xi}^0_{n+p}) - \triangledown_k f(Z,\boldsymbol{\xi}^0)\right| < C.\max(1,||X||).||\boldsymbol{\xi}^0_n - \boldsymbol{\xi}^0||_0 \\
&\text{if } ||\boldsymbol{\xi}^0_n - \boldsymbol{\xi}^0||_0 < d.
\end{aligned} \tag{14.11}
$$

14.2. Part (a) for $k > p$

Next, it will be shown that, without the need for Assumption 2.3,

$$\sup_{m\in\mathbb{N}} \sup_{\boldsymbol{\delta}\in\Delta_0(M)} \sup_{0\leq u\leq 1} |\psi_m(u|\boldsymbol{\delta})| < \infty \tag{14.12}$$

and that there exists a constant $C > 0$ such that for all $m \in \mathbb{N}$,

$$\sup_{0\leq u\leq 1} |\psi_m(u|\boldsymbol{\delta}) - \psi_m(u|\boldsymbol{\delta}^0)| < C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 \tag{14.13}$$

It follows then straightforwardly from (14.3), (14.12) and (14.13), similar to (14.11), that there exist constants $C > 0$ and $d > 0$ such that

$$\sup_{m\in\mathbb{N}} \left|\bigtriangledown_{p+m} f(Z, \boldsymbol{\xi}_n^0) - \bigtriangledown_{p+m} f(Z, \boldsymbol{\xi}^0)\right| \leq C. \max(1, ||X||).||\boldsymbol{\xi}_n^0 - \boldsymbol{\xi}^0||_0$$
$$\text{if } ||\boldsymbol{\xi}_n^0 - \boldsymbol{\xi}^0||_0 < d. \tag{14.14}$$

To prove (14.12), observe from the easy equality

$$\bigtriangledown_m h(u|\boldsymbol{\delta}) = 2\frac{\left(1 + \sum_{k=1}^{\infty} \delta_k \sqrt{2}\cos(k\pi u)\right)\sqrt{2}\cos(m\pi u)}{1 + \sum_{k=1}^{\infty} \delta_k^2} - 2\delta_m \frac{h(u|\boldsymbol{\delta})}{1 + \sum_{k=1}^{\infty} \delta_k^2}$$

that

$$\begin{aligned}
&\frac{1}{2}\left(1 + \sum_{k=1}^{\infty} \delta_k^2\right) \bigtriangledown_m H(u|\boldsymbol{\delta}) \\
&= \sqrt{2}\frac{\sin(m\pi u)}{m\pi} + \sum_{k=1}^{\infty} \delta_k \frac{\sin((m+k)\pi u)}{(m+k)\pi} \\
&+ \sum_{k=1}^{\infty} \delta_k I(k \neq m)\frac{\sin((m-k)\pi u)}{(m-k)\pi} + \delta_m (u - H(u|\boldsymbol{\delta})) \\
&= -(-1)^m \sqrt{2}\frac{\sin(m\pi(1-u))}{m\pi} - \sum_{k=1}^{\infty} \delta_k (-1)^{m+k} \frac{\sin((m+k)\pi(1-u))}{(m+k)\pi} \\
&- \sum_{k=1}^{\infty} \delta_k I(k \neq m)(-1)^{m-k} \frac{\sin((m-k)\pi(1-u))}{(m-k)\pi} \\
&+ \delta_m (u - H(u|\boldsymbol{\delta}))
\end{aligned} \tag{14.15}$$

hence

$$\frac{\nabla_m H(u|\boldsymbol{\delta})}{u} = \frac{2}{1+\sum_{k=1}^{\infty}\delta_k^2}\left\{\sqrt{2}\frac{\sin(m\pi u)}{m\pi u} + \sum_{k=1}^{\infty}\delta_k\frac{\sin((m+k)\pi u)}{(m+k)\pi u}\right.$$
$$\left. + \sum_{k=1}^{\infty}\delta_k I(k\neq m)\frac{\sin((m-k)\pi u)}{(m-k)\pi u} + \delta_m - \delta_m\frac{H(u|\boldsymbol{\delta})}{u}\right\} \tag{14.16}$$

and

$$\frac{\nabla_m H(u|\boldsymbol{\delta})}{1-u} = \frac{2}{1+\sum_{k=1}^{\infty}\delta_k^2}\left\{-(-1)^m\sqrt{2}\frac{\sin(m\pi(1-u))}{m\pi(1-u)}\right.$$
$$- \sum_{k=1}^{\infty}\delta_k(-1)^{m+k}\frac{\sin((m+k)\pi(1-u))}{(m+k)\pi(1-u)}$$
$$- \sum_{k=1}^{\infty}\delta_k I(k\neq m)(-1)^{m-k}\frac{\sin((m-k)\pi(1-u))}{(m-k)\pi(1-u)}$$
$$\left. - \delta_m + \delta_m\left(\frac{1-H(u|\boldsymbol{\delta})}{1-u}\right)\right\} \tag{14.17}$$

Adding these two expressions up and taking absolute values yield

$$\begin{aligned} |\psi_m(u|\boldsymbol{\delta})| &= \left|\frac{\nabla_m H(u|\boldsymbol{\delta})}{u} + \frac{\nabla_m H(u|\boldsymbol{\delta})}{1-u}\right| \\ &\leq 2\frac{2\sqrt{2}+4\sum_{k=1}^{\infty}|\delta_k| + |\delta_m|\left(H(u|\boldsymbol{\delta})/u + (1-H(u|\boldsymbol{\delta}))/(1-u)\right)}{1+\sum_{k=1}^{\infty}\delta_k^2} \\ &\leq 2\frac{2\sqrt{2}+4\sum_{k=1}^{\infty}|\delta_k| + 2|\delta_m|\sup_{0\leq u\leq 1}h(u|\boldsymbol{\delta})}{1+\sum_{k=1}^{\infty}\delta_k^2}, \end{aligned}$$

which implies (14.12).

Similarly, it is not hard to verify from (14.15) that there exists a constant $C>0$ such that (14.13) holds for all $m\in\mathbb{N}$.

14.3. Part (b)

Part (b) of Lemma 7.1 follows trivially from (14.1), (14.3) and

$$E[Y|X] = H(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0) \text{ a.s.}$$

14.4. Part (c)

It follows from (14.1) and (14.3) that in general

$$\max_{1\le k\le p} E\left[\left(\triangledown_k f(Z,\boldsymbol{\xi}^0)\right)^2\right] \;\le\; \sup_{0\le u\le 1} h\left(u|\boldsymbol{\delta}^0\right)^2 E\left[\phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)^2 \max(1,||X||^2)\right],$$

$$\sup_{m\in\mathbb{N}} E\left[\left(\triangledown_{p+m} f(Z,\boldsymbol{\xi}^0)\right)^2\right] \;\le\; \sup_{m\in\mathbb{N}} \sup_{\boldsymbol{\delta}\in\Delta_0(M)} \sup_{0\le u\le 1} |\psi_m(u|\boldsymbol{\delta})| E\left[\phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)^2\right],$$

so that by Assumption 2.3 and its implication (14.8),

$$\sup_{k\in\mathbb{N}} E\left[\left(\triangledown_k f(Z,\boldsymbol{\xi}^0)\right)^2\right] \le C.E\left[\max(1,||X||^2)\right] < \infty$$

for some constant $C > 0$.

15. Proof of Lemma 7.2

15.1. Part (a)

The second derivatives $\triangledown_{k,m} f(Z,\boldsymbol{\xi})$ for $k,m = 1,2,...,p$ are

$$\begin{aligned}
\triangledown_{k,m} f(Z,\boldsymbol{\xi}) &= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) \\
&\times \frac{G''(\widetilde{X}'\theta)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)} \widetilde{X}_k \widetilde{X}_m \\
&- \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)^2 \\
&\times \left(\frac{G'(\widetilde{X}'\theta)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)}\right)^2 \\
&\times \left(1 - 2H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) \widetilde{X}_k \widetilde{X}_m \\
&+ \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h^{(1)}\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) \\
&\times \frac{G'(\widetilde{X}'\theta)^2}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)} \widetilde{X}_k \widetilde{X}_m \\
&- h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)^2 \frac{G'(\widetilde{X}'\theta)^2}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)} \widetilde{X}_k \widetilde{X}_m.
\end{aligned}$$

Using the function $\phi(u|\boldsymbol{\delta})$ in (14.2) and the fact that for the logistic distribution function $G(x)$, $G'(x) = G(x)(1 - G(x))$ and $G''(x) = G(x)(1 - G(x))(1 - 2G(x))$, we can write

$$\begin{aligned}
&\nabla_{k,m} f(Z, \boldsymbol{\xi}) \\
&= \Big\{ \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) (1 - 2G(\widetilde{X}'\theta))\phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) \\
&- \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)^2 \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta})^2 \left(1 - 2H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) \\
&+ \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) h^{(1)}\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) G'(\widetilde{X}'\theta) \\
&-h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)^2 \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) G'(\widetilde{X}'\theta) \Big\} \widetilde{X}_k \widetilde{X}_m, \ k, m = 1, ..., p \qquad (15.1)
\end{aligned}$$

Recall from (14.10) that there exist constants $C > 0$ and $d > 0$ such that

$$\sup_{\theta \in \Theta} \left| \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) - \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0) \right| \leq C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 \text{ if } ||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 < d.$$

Moreover, note that

$$u(1 - u)\phi'(u|\boldsymbol{\delta}^0) = (1 - 2u)\phi(u|\boldsymbol{\delta}^0) - h(u|\boldsymbol{\delta}^0)(1 - 2H(u|\boldsymbol{\delta}^0)).\phi(u|\boldsymbol{\delta}^0)^2,$$

so that

$$\begin{aligned}
|\mathrm{d}\phi(G(x)|\boldsymbol{\delta}^0)/\mathrm{d}x| &= G(x)(1 - G(x))|\phi'(G(x)|\boldsymbol{\delta}^0)| \\
&\leq \phi(G(x)|\boldsymbol{\delta}^0) + h(G(x)|\boldsymbol{\delta}^0)\phi(G(x)|\boldsymbol{\delta}^0)^2.
\end{aligned}$$

Therefore, by (14.8) and the mean value theorem, with mean value θ_*,

$$\begin{aligned}
&\left| \phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0) - \phi(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0) \right| \\
&\quad \leq \left(\phi(G(\widetilde{X}'\theta_*)|\boldsymbol{\delta}^0) + \sup_{0 \leq u \leq 1} h(u|\boldsymbol{\delta}^0).\phi(G(\widetilde{X}'\theta_*)|\boldsymbol{\delta}^0)^2 \right) \\
&\qquad \times \max(1, ||X||).||\theta - \theta_0|| \\
&\quad \leq C.\max(1, ||X||).||\theta - \theta_0|| \qquad (15.2)
\end{aligned}$$

for some constant C.

Furthermore, by Lemma 5.1 and the mean value theorem, with C a generic constant,

$$\begin{aligned}
\left|G(\widetilde{X}'\theta) - G(\widetilde{X}'\theta_0)\right| &\leq C.\max(1, ||X||).||\theta - \theta_0|| \\
\left|G'(\widetilde{X}'\theta) - G'(\widetilde{X}'\theta_0)\right| &\leq C.\max(1, ||X||).||\theta - \theta_0|| \\
\left|h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) - h\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)\right| &\leq C.\max(1, ||X||).||\theta - \theta_0|| + C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 \\
\left|h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)^2 - h\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)^2\right| &\leq C.\max(1, ||X||).||\theta - \theta_0|| + C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 \\
\left|H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) - H\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)\right| &\leq C.\max(1, ||X||).||\theta - \theta_0|| + C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 \\
\left|h^{(1)}\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) - h^{(1)}\left(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0\right)\right| &\leq C.\max(1, ||X||).||\theta - \theta_0|| + C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_1
\end{aligned}$$

where the latter follows from

$$\sup_{\boldsymbol{\delta}\in\Delta_2(M)} \sup_{0\leq u\leq 1} h^{(2)}(u|\boldsymbol{\delta}) < \infty \tag{15.3}$$

due to Lemma 5.1 and the choice $\ell = 2$ in Assumption 7.1,[3] and

$$\sup_{0\leq u\leq 1} |h^{(1)}(u|\boldsymbol{\delta}) - h^{(1)}(u|\boldsymbol{\delta}^0)| \leq C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_1.$$

It follows now easily that for $k, m = 1, 2, .., p$, and some constant $C > 0$,

$$| \bigtriangledown_{k,m} f(Z, \boldsymbol{\xi}) - \bigtriangledown_{k,m} f(Z, \boldsymbol{\xi}^0)| \leq C.\max(1, ||X||^3).||\boldsymbol{\xi} - \boldsymbol{\xi}^0||_1 \tag{15.4}$$

if $||\boldsymbol{\xi} - \boldsymbol{\xi}^0||_1$ is sufficiently small.

Next, observe that for $k = 1, ..., p$ and $m \in \mathbb{N}$,

$$\begin{aligned}
\bigtriangledown_{k,p+m} f(Z, \boldsymbol{\xi}) &= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) \bigtriangledown_m h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) \\
&\quad \times \frac{G'(\widetilde{X}'\theta)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1 - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)} \widetilde{X}_k \\
&\quad - \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right) \bigtriangledown_m H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)
\end{aligned}$$

[3] The inequality (15.3) is one of the reasons why $\ell = 2$ is needed in Assumption 7.1.

$$\times \frac{G'(\widetilde{X}'\theta)\left(1-2H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)^2\left(1-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)^2}\widetilde{X}_k$$
$$-h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right) G'(\widetilde{X}'\theta) \bigtriangledown_m H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)$$
$$\times \frac{1}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)}\widetilde{X}_k,$$

which can be written as

$$\begin{aligned}
&\bigtriangledown_{k,p+m} f(Z,\boldsymbol{\xi}) \\
&= \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta})G'(\widetilde{X}'\theta)\widetilde{X}_k \\
&- \left(Y - H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta})^2 \\
&\times \left(1-2H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\widetilde{X}_k \\
&-h\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\phi(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\widetilde{X}_k, \qquad (15.5)
\end{aligned}$$

where $\psi_m(u|\boldsymbol{\delta})$ is defined in (14.4).

To prove that for $k = 1, 2, ..., p$ and $m \in \mathbb{N}$,

$$|\bigtriangledown_{k,m+p} f(Z,\boldsymbol{\xi}) - \bigtriangledown_{k,m+p} f(Z,\boldsymbol{\xi}^0)| \leq C.\max(1, ||X||^2).||\boldsymbol{\xi} - \boldsymbol{\xi}^0||_0 \qquad (15.6)$$

if $||\boldsymbol{\xi} - \boldsymbol{\xi}^0||_1$ is sufficiently small, it suffices to show that for some constant C,

$$\begin{aligned}
|\psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) - \psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0)| &\leq C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0 \qquad (15.7)\\
|\psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0) - \psi_m(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)| &\leq C.\max(1, ||X||).||\theta - \theta_0|| \qquad (15.8)
\end{aligned}$$

as all the other inequalities involved have already been derived.

It follows straightforwardly from (14.16) and (14.17) that

$$\sup_{m\in\mathbb{N}} \sup_{0\leq u\leq 1} \frac{|\bigtriangledown_m H(u|\boldsymbol{\delta}) - \bigtriangledown_m H(u|\boldsymbol{\delta}^0)|}{u} \leq C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0,$$
$$\sup_{m\in\mathbb{N}} \sup_{0\leq u\leq 1} u\left|\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\bigtriangledown_m H(u|\boldsymbol{\delta}^0)}{u}\right)\right| \leq C.||\boldsymbol{\delta}^0||_1 < \infty,$$
$$\sup_{m\in\mathbb{N}} \sup_{0\leq u\leq 1} \frac{|\bigtriangledown_m H(u|\boldsymbol{\delta}) - \bigtriangledown_m H(u|\boldsymbol{\delta}^0)|}{1-u} \leq C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0,$$
$$\sup_{m\in\mathbb{N}} \sup_{0\leq u\leq 1} (1-u)\left|\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\bigtriangledown_m H(u|\boldsymbol{\delta}^0)}{1-u}\right)\right| \leq C.||\boldsymbol{\delta}^0||_1 < \infty,$$

hence

$$\sup_{m\in\mathbb{N}}\sup_{0\le u\le 1}|\psi_m(u|\boldsymbol{\delta})-\psi_m(u|\boldsymbol{\delta}^0)| \le C.||\boldsymbol{\delta}-\boldsymbol{\delta}^0||_0, \qquad (15.9)$$

$$\sup_{m\in\mathbb{N}}\sup_{0\le u\le 1} u(1-u)\psi'_m(u|\boldsymbol{\delta}^0) < \infty. \qquad (15.10)$$

Clearly, (15.9) implies (15.7), and by the mean value theorem, (15.10) implies (15.8).

Finally, for $k,m\in\mathbb{N}$ we have

$$\begin{aligned}\bigtriangledown_{p+k,p+m}f(Z,\boldsymbol{\xi}) &= \left(Y-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\\ &\times\frac{\bigtriangledown_{k,m}H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)}\\ &-\left(Y-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\left(1-2H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\\ &\times\frac{\bigtriangledown_k H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\bigtriangledown_m H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)^2\left(1-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)^2}\\ &-\frac{\bigtriangledown_k H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\bigtriangledown_m H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)}{H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)},\end{aligned}$$

which can be written as

$$\begin{aligned}&\bigtriangledown_{p+k,p+m}f(Z,\boldsymbol{\xi})\\ &=\left(Y-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\psi_{m,k}(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\\ &-\left(Y-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\left(1-2H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\psi_k(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\\ &-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\left(1-H\left(G(\widetilde{X}'\theta)|\boldsymbol{\delta}\right)\right)\psi_k(G(\widetilde{X}'\theta)|\boldsymbol{\delta})\psi_m(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) \qquad (15.11)\end{aligned}$$

where

$$\psi_{m,k}(u|\boldsymbol{\delta})=\frac{\bigtriangledown_{k,m}H(u|\boldsymbol{\delta})}{u(1-u)}.$$

I will show that

$$\sup_{k,m\in\mathbb{N}}\sup_{\boldsymbol{\delta}\in\Delta_1(M)}\sup_{0\le u\le 1}|\psi_{k,m}(u|\boldsymbol{\delta})|<\infty, \qquad (15.12)$$

$$\sup_{k,m\in\mathbb{N}} \sup_{0\le u\le 1} |\psi_{k,m}(u|\boldsymbol{\delta}) - \psi_{k,m}(u|\boldsymbol{\delta}^0)| \le C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0, \tag{15.13}$$

$$u(1-u)\psi'_{k,m}(u|\boldsymbol{\delta}) = \triangledown_{k,m} h(u|\boldsymbol{\delta}) - (1-2u)\psi_{k,m}(u|\boldsymbol{\delta}), \tag{15.14}$$

so that

$$|\psi_{m,k}(G(\widetilde{X}'\theta)|\boldsymbol{\delta}) - \psi_{m,k}(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0)| \le C.||\boldsymbol{\delta} - \boldsymbol{\delta}^0||_0$$

and by the mean value theorem,

$$\left|\psi_{m,k}(G(\widetilde{X}'\theta)|\boldsymbol{\delta}^0) - \psi_{m,k}(G(\widetilde{X}'\theta_0)|\boldsymbol{\delta}^0)\right| \le C.\max(1, ||X||).||\theta - \theta_0||.$$

It follows then easily that

$$| \triangledown_{p+k,p+m} f(Z, \boldsymbol{\xi}) - \triangledown_{p+k,p+m} f(Z, \boldsymbol{\xi}^0)| \le C.\max(1, ||X||).||\boldsymbol{\xi} - \boldsymbol{\xi}^0||_0. \tag{15.15}$$

To prove (15.12)-(15.14), observe from (14.15) that for $k \neq m$,

$$\begin{aligned}
&\delta_k \triangledown_m H(u|\boldsymbol{\delta}) + \frac{1}{2}\left(1 + \sum_{k=1}^{\infty} \delta_k^2\right) \triangledown_{k,m} H(u|\boldsymbol{\delta}) \\
&\qquad = 2\frac{\sin((m+k)\pi u)}{(m+k)\pi} - \delta_m \left(\triangledown_k H(u|\boldsymbol{\delta})\right) \\
&\qquad = -2(-1)^{m+k}\frac{\sin((m+k)\pi(1-u))}{(m+k)\pi} - \delta_m \triangledown_k H(u|\boldsymbol{\delta})
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{\triangledown_{k,m} H(u|\boldsymbol{\delta})}{u} &= \frac{2}{1 + \sum_{i=1}^{\infty} \delta_i^2} \\
&\quad \times \left\{2\frac{\sin((m+k)\pi u)}{(m+k)\pi u} - (1-u)\delta_m \left(\psi_m(u|\boldsymbol{\delta}) + \psi_k(u|\boldsymbol{\delta})\right)\right\}
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\triangledown_{k,m} H(u|\boldsymbol{\delta})}{1-u} = \frac{2}{1 + \sum_{i=1}^{\infty} \delta_i^2} \\
&\quad \times \left\{-2(-1)^{m+k}\frac{\sin((m+k)\pi(1-u))}{(m+k)\pi(1-u)} - u.\delta_m \left(\psi_m(u|\boldsymbol{\delta}) + \psi_k(u|\boldsymbol{\delta})\right)\right\}
\end{aligned}$$

so that

$$\psi_{m,k}(u|\boldsymbol{\delta}) = \frac{\nabla_{k,m}H(u|\boldsymbol{\delta})}{u(1-u)} = \frac{-2\delta_m\left(\psi_m(u|\boldsymbol{\delta}) + \psi_k(u|\boldsymbol{\delta})\right)}{1+\sum_{i=1}^{\infty}\delta_i^2} \tag{15.16}$$

$$+\frac{2}{1+\sum_{i=1}^{\infty}\delta_i^2}\left\{2\frac{\sin((m+k)\pi u)}{(m+k)\pi u} - 2(-1)^{m+k}\frac{\sin((m+k)\pi(1-u))}{(m+k)\pi(1-u)}\right\}$$

Similarly, for $k = m$ we have

$$\frac{\nabla_{m,m}H(u|\boldsymbol{\delta})}{u} = \frac{2}{1+\sum_{i=1}^{\infty}\delta_i^2} \times\left\{\frac{\sin(2m\pi u)}{2m\pi u} + 1 - \frac{H(u|\boldsymbol{\delta})}{u} - 2\delta_m\frac{\nabla_m H(u|\boldsymbol{\delta})}{u}\right\}$$

and

$$\frac{\nabla_{m,m}H(u|\boldsymbol{\delta})}{1-u} = \frac{2}{1+\sum_{i=1}^{\infty}\delta_i^2} \times\left\{-\frac{\sin(2m\pi(1-u))}{2m\pi(1-u)} - 1 + \frac{1-H(u|\boldsymbol{\delta})}{1-u} - 2\delta_m\frac{\nabla_m H(u|\boldsymbol{\delta})}{1-u}\right\}$$

hence

$$\psi_{m,m}(u|\boldsymbol{\delta}) = \frac{\nabla_{m,m}H(u|\boldsymbol{\delta})}{u(1-u)} = \frac{2}{1+\sum_{i=1}^{\infty}\delta_i^2} \tag{15.17}$$

$$\times\left\{\frac{\sin(2m\pi u)}{2m\pi u} - \frac{\sin(2m\pi(1-u))}{2m\pi(1-u)} + \frac{1-H(u|\boldsymbol{\delta})}{1-u} - \frac{H(u|\boldsymbol{\delta})}{u} - 2\delta_m\psi_m(u|\boldsymbol{\delta})\right\}$$

The results (15.12)-(15.14) follow now straightforwardly from (15.16) and (15.17).

Combining (15.4), (15.6) and (15.15), part (a) of Lemma 7.2 follows.

15.2. Part (b)

Part (b) of Lemma 7.2 follows trivially from (15.1), (15.5) and (15.11).

15.3. Part (c)

Using the function $\phi(u|\boldsymbol{\delta}^0)$ in (14.2), the notation (14.6), and the fact that for the logistic distribution function $G(x)$, $G'(x) = G(x)(1-G(x))$ and $G''(x) = G(x)(1-G(x))(1-2G(x))$, it is not hard to verify that for $k, m = 1, 2, ..., p$,

$$E[\nabla_{k,m}f(Z,\boldsymbol{\xi}^0)|X] = -\frac{h\left(U_X|\boldsymbol{\delta}^0\right)^2\left(U_X(1-U_X)\right)^2}{H\left(U_X|\boldsymbol{\delta}^0\right)\left(1-H\left(U_X|\boldsymbol{\delta}^0\right)\right)}\widetilde{X}_k\widetilde{X}_m$$

Hence, the matrix $B_{p,p}$ takes the form $B_{p,p} = -\overline{B}_p$ where

$$\overline{B}_p = E\left[\frac{h\left(U_X|\boldsymbol{\delta}^0\right)^2 (U_X(1-U_X))^2}{H\left(U_X|\boldsymbol{\delta}^0\right)\left(1-H\left(U_X|\boldsymbol{\delta}^0\right)\right)}\begin{pmatrix} 1 & X' \\ X & XX' \end{pmatrix}\right].$$

which is a.s. finite.

To prove that $\overline{B}_p$ is nonsingular, suppose that there exists a nonzero vector $\gamma = (\gamma_0, \gamma_1')' \in \mathbb{R} \times \mathbb{R}^{p-1}$ such that $\gamma'\overline{B}_p\gamma = 0$. Obviously, this is only possible if $\gamma_0^2 + 2\gamma_1'X + (\gamma_1'X)^2 = 0$ a.s.,[4] hence $\gamma_1'X = -\gamma_0$ a.s. and thus $\gamma_1'\mathrm{Var}(X)\gamma_1 = 0$. Since by part (c) of Assumption 2.1, $\mathrm{Var}(X)$ is nonsingular, it follows that $\gamma_1 = 0$, hence $\gamma_0 = 0$, which contradicts $\gamma \neq 0$. Thus, $\overline{B}_p$ is nonsingular.

Recall from (14.1) that for $k = 1, 2, .., p$,

$$\triangledown_k f(Z, \boldsymbol{\xi}^0) = \left(Y - H\left(U_X|\boldsymbol{\delta}^0\right)\right) h\left(U_X|\boldsymbol{\delta}^0\right)\phi(U_X|\boldsymbol{\delta}^0)\widetilde{X}_k,$$

hence for $k, m = 1, 2, ..., p$,

$$\begin{aligned} &E\left[\triangledown_k f(Z, \boldsymbol{\xi}^0)\triangledown_m f(Z, \boldsymbol{\xi}^0)|X\right] \\ &= E\left[\left(Y - H\left(U_X|\boldsymbol{\delta}^0\right)\right)^2\Big| X\right] h\left(U_X|\boldsymbol{\delta}^0\right)^2\phi(U_X|\boldsymbol{\delta}^0)^2\widetilde{X}_k\widetilde{X}_m \\ &= H\left(U_X|\boldsymbol{\delta}^0\right)\left(1 - H\left(U_X|\boldsymbol{\delta}^0\right)\right) h\left(U_X|\boldsymbol{\delta}^0\right)^2\phi(U_X|\boldsymbol{\delta}^0)^2\widetilde{X}_k\widetilde{X}_m \\ &= \frac{h\left(U_X|\boldsymbol{\delta}^0\right)^2 (U_X(1-U_X))^2}{H\left(U_X|\boldsymbol{\delta}^0\right)\left(1-H\left(U_X|\boldsymbol{\delta}^0\right)\right)}\widetilde{X}_k\widetilde{X}_m \end{aligned}$$

so that $V_p = -B_{p.p}$, where V_p is the variance matrix of $(\triangledown_1 f(Z, \boldsymbol{\xi}^0), ..., \triangledown_p f(Z, \boldsymbol{\xi}^0))'$.

Also, it is not hard to verify that for $k = 1, ..., p$ and $m \in \mathbb{N}$,

$$E[\triangledown_{k,p+m} f(Z, \boldsymbol{\xi}^0)|X] = -h\left(U_X|\boldsymbol{\delta}^0\right)\psi_m(U_X|\boldsymbol{\delta}^0)\phi(U_X|\boldsymbol{\delta}^0)\widetilde{X}_k, \tag{15.18}$$

and for $k, m \in \mathbb{N}$,

$$\begin{aligned} E[\triangledown_{p+k,p+m} f(Z, \boldsymbol{\xi}^0)|X] &= -H\left(U_X|\boldsymbol{\delta}^0\right)\left(1 - H\left(U_X|\boldsymbol{\delta}^0\right)\right) \\ &\quad \times\psi_k(U_X|\boldsymbol{\delta}^0)\psi_m(U_X)|\boldsymbol{\delta}^0). \end{aligned} \tag{15.19}$$

whereas by (14.1) and (14.3),

$$E\left[\triangledown_k f(Z, \boldsymbol{\xi}^0)\triangledown_{p+m} f(Z, \boldsymbol{\xi}^0)|X\right] = h\left(U_X|\boldsymbol{\delta}^0\right)\psi_m(U_X|\boldsymbol{\delta}^0)\phi(U_X|\boldsymbol{\delta}^0)\widetilde{X}_k$$

[4] See for example Lemma 3.1 in Bierens (2004).

for $k = 1, ..., p$ and $m \in \mathbb{N}$, and

$$\begin{aligned} E\left[\bigtriangledown_{p+k} f(Z, \boldsymbol{\xi}^0) \bigtriangledown_{p+m} f(Z, \boldsymbol{\xi}^0) | X\right] &= H\left(U_X | \boldsymbol{\delta}^0\right)\left(1 - H\left(U_X | \boldsymbol{\delta}^0\right)\right) \\ &\quad \times \psi_k(U_X | \boldsymbol{\delta}^0) \psi_m(U_X | \boldsymbol{\delta}^0) \end{aligned}$$

for $k, m \in \mathbb{N}$. Hence,

$$V_n = -B_{n,n} \text{ for all } n \in \mathbf{N},$$

where V_n is defined in Theorem 6.2.

To show that V_n is nonsingular, suppose first that V_{p+1} is singular. Since it has already been verified that V_p is nonsingular, singularity of V_{p+1} implies that $\bigtriangledown_{p+1} f(Z, \boldsymbol{\xi}^0)$ is a.s. a linear combination of $(\bigtriangledown_1 f(Z, \boldsymbol{\xi}^0), ..., \bigtriangledown_p f(Z, \boldsymbol{\xi}^0))'$:

$$\begin{aligned} &\left(Y - H\left(U_X | \boldsymbol{\delta}^0\right)\right) \psi_1(U_X | \boldsymbol{\delta}^0) \phi(U_X | \boldsymbol{\delta}^0) \\ &\quad = \left(Y - H\left(U_X | \boldsymbol{\delta}^0\right)\right) h\left(U_X | \boldsymbol{\delta}^0\right) \phi(U_X | \boldsymbol{\delta}^0) \widetilde{X}' \gamma \text{ a.s.} \end{aligned}$$

for some $\gamma \in \mathbb{R}^p$, hence

$$\frac{\bigtriangledown_1 H(U_X | \boldsymbol{\delta}^0)}{U_X(1 - U_X) h(U_X | \boldsymbol{\delta}^0)} = \frac{\psi_1(U_X | \boldsymbol{\delta}^0)}{h(U_X | \boldsymbol{\delta}^0)} = \widetilde{X}' \gamma \text{ a.s.}$$

Next, write $U_X = G(\alpha_0 + \beta_{0,1} X_1 + \beta_{0,2}' X_2)$ and $\widetilde{X}' \gamma = \gamma_1 + \gamma_2 X_1 + \gamma_3' X_2$, and recall from Assumption 2.1 that X_1 has support $\mathbb{R}$, conditional on X_2. Therefore, we may take the partial derivatives to X_1:

$$\partial U_X / \partial X_1 = U_X(1 - U_X) \beta_{0,1}, \ \partial \widetilde{X}' \gamma / \partial X_1 = \gamma_2.$$

Then

$$\begin{aligned} \frac{\gamma_2}{\beta_{0,1}} &= -\frac{\bigtriangledown_1 H(U_X | \boldsymbol{\delta}^0)}{\left(U_X(1 - U_X) h(U_X | \boldsymbol{\delta}^0)\right)^2} \\ &\quad \times \left((1 - 2U_X) h(U_X | \boldsymbol{\delta}^0) + U_X(1 - U_X) h^{(1)}(U_X | \boldsymbol{\delta}^0)\right) \\ &\quad \times U_X(1 - U_X) + \frac{\bigtriangledown_1 h(U_X | \boldsymbol{\delta}^0)}{U_X(1 - U_X) h(U_X | \boldsymbol{\delta}^0)} U_X(1 - U_X) \\ &= \frac{\bigtriangledown_1 h(U_X | \boldsymbol{\delta}^0)}{h(U_X | \boldsymbol{\delta}^0)} - \frac{\bigtriangledown_1 H(U_X | \boldsymbol{\delta}^0)(1 - 2U_X)}{U_X(1 - U_X) h(U_X | \boldsymbol{\delta}^0)} \\ &\quad - \frac{h^{(1)}(U_X | \boldsymbol{\delta}^0) \bigtriangledown_1 H(U_X | \boldsymbol{\delta}^0)}{h(U_X | \boldsymbol{\delta}^0)^2} \end{aligned}$$

Thus, the singularity of V_{p+1} implies that for all $u \in [0,1]$,

$$\frac{\gamma_2}{\beta_{0,1}} \equiv \frac{\nabla_1 h(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)} - \frac{\nabla_1 H(u|\boldsymbol{\delta}^0)\,(1-2u)}{u(1-u)h(u|\boldsymbol{\delta}^0)} - \frac{h^{(1)}(u|\boldsymbol{\delta}^0)\,\nabla_1\, H(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)^2}$$

Taking the derivative to u yields:

$$\begin{aligned}
0 \quad \equiv \quad & \frac{\nabla_1 h^{(1)}(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)} - \frac{\nabla_1 h(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)^2} h^{(1)}(u|\boldsymbol{\delta}^0) \\
& - \frac{\nabla_1 h(u|\boldsymbol{\delta}^0)\,(1-2u)}{u(1-u)h(u|\boldsymbol{\delta}^0)} + 2\frac{\nabla_1 H(u|\boldsymbol{\delta}^0)}{u(1-u)h(u|\boldsymbol{\delta}^0)} \\
& + \frac{\nabla_1 H(u|\boldsymbol{\delta}^0)\,(1-2u)}{(u(1-u))^2 h(u|\boldsymbol{\delta}^0)^2}\,(1-2u)\,h(u|\boldsymbol{\delta}^0) \\
& + \frac{\nabla_1 H(u|\boldsymbol{\delta}^0)\,(1-2u)}{(u(1-u))^2 h(u|\boldsymbol{\delta}^0)^2} u(1-u) h^{(1)}(u|\boldsymbol{\delta}^0) \\
& - \frac{h^{(2)}(u|\boldsymbol{\delta}^0)\,\nabla_1\, H(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)^2} - \frac{h^{(1)}(u|\boldsymbol{\delta}^0)\,\nabla_1\, h(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)^2} \\
& + 2\frac{h^{(1)}(u|\boldsymbol{\delta}^0)\,\nabla_1\, H(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)^3} h^{(1)}(u|\boldsymbol{\delta}^0)
\end{aligned}$$

hence

$$\begin{aligned}
0 \quad \equiv \quad & \nabla_1 h^{(1)}(u|\boldsymbol{\delta}^0) - \frac{\nabla_1 h(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)} h^{(1)}(u|\boldsymbol{\delta}^0) \\
& - \frac{\nabla_1 h(u|\boldsymbol{\delta}^0)\,(1-2u)}{u(1-u)} + 2\frac{\nabla_1 H(u|\boldsymbol{\delta}^0)}{u(1-u)} \\
& + \frac{\nabla_1 H(u|\boldsymbol{\delta}^0)\,(1-2u)^2}{(u(1-u))^2} + \frac{\nabla_1 H(u|\boldsymbol{\delta}^0)\,(1-2u)}{u(1-u)h(u|\boldsymbol{\delta}^0)} h^{(1)}(u|\boldsymbol{\delta}^0) \\
& - \frac{h^{(2)}(u|\boldsymbol{\delta}^0)\,\nabla_1\, H(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)} - \frac{h^{(1)}(u|\boldsymbol{\delta}^0)\,\nabla_1\, h(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)} \\
& + 2\frac{h^{(1)}(u|\boldsymbol{\delta}^0)^2\,\nabla_1\, H(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)^2}
\end{aligned} \tag{15.20}$$

Now observe from (14.16) and (14.17) that

$$\lim_{u \downarrow 0} \frac{\nabla_m H(u|\boldsymbol{\delta}^0)}{u(1-u)} = 2\sqrt{2}\frac{\sqrt{h(0|\boldsymbol{\delta}^0)}}{\sqrt{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}} - \sqrt{2}\delta_{0,m}\frac{h(0|\boldsymbol{\delta}^0)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}$$

whereas by Theorem 3.1,

$$
\begin{aligned}
\nabla_m h(u|\boldsymbol{\delta}^0) &= \frac{2\left(1+\sum_{k=1}^{\infty}\delta_{0,k}\sqrt{2}\cos(k\pi u)\right)\sqrt{2}\cos(m\pi u)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2} \\
&\quad -2\frac{h(u|\boldsymbol{\delta}^0).\delta_{0,m}}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2} \\
h^{(1)}(u|\boldsymbol{\delta}^0) &= -2\frac{\left(1+\sum_{k=1}^{\infty}\delta_{0,k}\sqrt{2}\cos(k\pi u)\right)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}\pi\sum_{k=1}^{\infty}k\delta_{0,k}\sqrt{2}\sin(k\pi u) \\
\nabla_m h^{(1)}(u|\boldsymbol{\delta}^0) &= \frac{h^{(1)}(u|\boldsymbol{\delta}^0)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}2\delta_{0,m} \\
&\quad -2\sqrt{2}\cos(m\pi u)\frac{\pi\sum_{k=1}^{\infty}k\delta_{0,k}\sqrt{2}\sin(k\pi u)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2} \\
&\quad -2\frac{\left(1+\sum_{k=1}^{\infty}\delta_{0,k}\sqrt{2}\cos(k\pi u)\right)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}\pi m\sqrt{2}\sin(m\pi u) \\
h^{(2)}(u|\boldsymbol{\delta}^0) &= -2\frac{\left(1+\sum_{k=1}^{\infty}\delta_{0,k}\sqrt{2}\cos(k\pi u)\right)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}\pi^2\sum_{k=1}^{\infty}k^2\delta_{0,k}\sqrt{2}\cos(k\pi u) \\
&\quad +2\frac{\left(\pi\sum_{k=1}^{\infty}k\delta_{0,k}\sqrt{2}\sin(k\pi u)\right)^2}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}
\end{aligned}
$$

(the latter result is the other reason why we need $\ell = 2$ in Assumption 7.1) so that

$$
\begin{aligned}
\lim_{u\downarrow 0}\nabla_m h(u|\boldsymbol{\delta}^0) &= 2\sqrt{2}\frac{\sqrt{h(0|\boldsymbol{\delta}^0)}}{\sqrt{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}} - 2\delta_{0,m}\frac{h(0|\boldsymbol{\delta}^0)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2} \\
\lim_{u\downarrow 0}h^{(1)}(u|\boldsymbol{\delta}^0) &= 0 \\
\lim_{u\downarrow 0}\nabla_m h^{(1)}(u|\boldsymbol{\delta}^0) &= 0 \\
\lim_{u\downarrow 0}h^{(2)}(u|\boldsymbol{\delta}^0) &= 0
\end{aligned}
$$

Letting $u \to 0$ in (15.20) it follows now that

$$
0 = 2\lim_{u\downarrow 0}\frac{\nabla_1 H(u|\boldsymbol{\delta}^0)}{u(1-u)} + \lim_{u\downarrow 0}\left(\frac{\nabla_1 H(u|\boldsymbol{\delta}^0)}{u(1-u)} - \nabla_1 h(u|\boldsymbol{\delta}^0)\right)\times\lim_{u\downarrow 0}\frac{1}{u(1-u)}
$$

$$
\begin{aligned}
&= 4\sqrt{2}\frac{\sqrt{h(0|\boldsymbol{\delta}^0)}}{\sqrt{1+\sum_{k=1}^{\infty}\delta_{0,k}^2}} - 2\sqrt{2}\delta_{0,1}\frac{h(0|\boldsymbol{\delta}^0)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2} \\
&\quad + \left(2-\sqrt{2}\right)\delta_{0,1}\frac{h(0|\boldsymbol{\delta}^0)}{1+\sum_{k=1}^{\infty}\delta_{0,k}^2} \times \lim_{u\downarrow 0}\frac{1}{u(1-u)},
\end{aligned}
$$

which is impossible. Consequently, V_{p+1} is nonsingular.

Along the same lines it can be shown that for all $m \in \mathbb{N}$ the variance matrix of $(\bigtriangledown_1 f(Z,\boldsymbol{\xi}^0), ..., \bigtriangledown_p f(Z,\boldsymbol{\xi}^0), \bigtriangledown_{p+m} f(Z,\boldsymbol{\xi}^0))'$ is nonsingular.

Next, suppose that for some $m \in \mathbb{N}$, V_{p+m} is nonsingular but V_{p+m+1} is singular, so that $\bigtriangledown_{p+m+1} f(Z,\boldsymbol{\xi}^0)$ is a.s. a linear combination of $(\bigtriangledown_1 f(Z,\boldsymbol{\xi}^0), ..., \bigtriangledown_{p+m} f(Z,\boldsymbol{\xi}^0))'$:

$$
\begin{aligned}
&\left(Y - H\left(U_X|\boldsymbol{\delta}^0\right)\right)\psi_{m+1}(U_X|\boldsymbol{\delta}^0)\phi(U_X|\boldsymbol{\delta}^0) \\
&\quad = \left(Y - H\left(U_X|\boldsymbol{\delta}^0\right)\right) h\left(U_X|\boldsymbol{\delta}^0\right)\phi(U_X|\boldsymbol{\delta}^0)\widetilde{X}'\gamma \\
&\quad + \sum_{k=1}^{m}\lambda_k\left(Y - H\left(U_X|\boldsymbol{\delta}^0\right)\right)\psi_k(U_X|\boldsymbol{\delta}^0)\phi(U_X|\boldsymbol{\delta}^0) \text{ a.s.}
\end{aligned}
$$

for some $\gamma \in \mathbb{R}^p$ and a nonzero vector $\lambda = (\lambda_1, ..., \lambda_m)' \in \mathbb{R}^m$, so that

$$
\psi_{m+1}(U_X|\boldsymbol{\delta}^0) - \sum_{k=1}^{m}\lambda_k.\psi_k(U_X|\boldsymbol{\delta}^0) = h\left(U_X|\boldsymbol{\delta}^0\right)\widetilde{X}'\gamma \text{ a.s.}
$$

Note that if $\lambda = 0$ then the variance matrix of $(\bigtriangledown_1 f(Z,\boldsymbol{\xi}^0), ..., \bigtriangledown_p f(Z,\boldsymbol{\xi}^0), \bigtriangledown_{p+m+1} f(Z,\boldsymbol{\xi}^0))'$ is singular, which is not true.

Replacing the operator $\bigtriangledown_1$ in (15.20) by $(\bigtriangledown_{m+1} - \sum_{k=1}^{m}\lambda_k\bigtriangledown_k)$ it follows now similar to the previous case that the singularity of V_{p+m+1} implies

$$
\begin{aligned}
\frac{\gamma_2}{\beta_{0,1}} &\equiv \frac{\bigtriangledown_{m+1}h(u|\boldsymbol{\delta}^0) - \sum_{k=1}^{m}\lambda_k\bigtriangledown_k h(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)} \\
&\quad - \left(\frac{1-2u}{u(1-u)} + \frac{h^{(1)}(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)}\right)\left(\frac{\bigtriangledown_{m+1}H(u|\boldsymbol{\delta}^0) - \sum_{k=1}^{m}\lambda_k\bigtriangledown_k H(u|\boldsymbol{\delta}^0)}{h(u|\boldsymbol{\delta}^0)}\right)
\end{aligned}
$$

which contradicts Assumption 7.2. Thus, V_{p+m+1} is nonsingular. By induction it follows now that V_n is nonsingular for all $n \in \mathbb{N}$, and so is $B_{n,n} = -V_n$.

15.4. Part (d)

Part (d) of Lemma 7.2 follows trivially from the nonsingularity of $B_{k+p,k+p}$.

References

Bierens, H.J. (2004) *Introduction to the Mathematical and Statistical Foundations of Econometrics*. Cambridge University Press.

Bierens, H.J. (2014) Consistency and asymptotic normality of sieve ML estimators under low-level conditions. Forthcoming in *Econometric Theory*.[5]

Billingsley, P. (1968) *Convergence of Probability Measures*. John Wiley.

van der Vaart, A.W. (1998) *Asymptotic Statistics*. Cambridge University Press.

van der Vaart, A.W. & J.A Wellner (1996) *Weak Convergence and Empirical Processes*. Springer.

[5]The accepted version of this paper is downloadable from http://econ.la.psu.edu/~hbierens/SNPMODELS.PDF