Online Supplemental Appendix for "Empirical Likelihood Ratio Tests of Conditional Moment Restrictions with Unknown Functions"

C Proof of Theorem 3.2.

Note that we still have

$$SCELR2 \;=\; = 2\sum_{i=1}^{n} \mathbf{1}_{in} \sum_{j=1}^{n} e_{ij} \left\{ 1 + \hat{\lambda}(x_i, \hat{\alpha}_n)' \rho(z_i, \hat{\alpha}_n) \right\} = \hat{D} + o_p(1)$$

under H_{1n}. Moreover, similar as the proof of Lemma E.3 and Lemma E.5, we have $b_n^{s/2}\hat{D}_1 = o_p(1)$. For the second term $\hat{D}_2$, by the same argument in the proof of Lemma E.4, we have $\hat{D}_2 = b_n^{-s}\{qR(K)vol(S_*) + o_p(1)\}$ under H_{1n}. Moreover, by the same argument in Lemma E.5, we have $b_n^{s/2}\hat{D}_3 = o_p(1)$.

Next, we focus on $\hat{D}_4$ term. The bandwidth $b_n = n^{-\alpha}$ for $0 < \alpha < \min \frac{1}{2s}(1 - \frac{4}{p})$ by Assumption A.3 (iii). We show that $b_n^{s/2}\hat{D}_4 \to N(\mu, 2qK^{**}vol(\mathcal{X}_n))$ under H_{1n}, where $\mu = \mathbb{E}[1\{x \in S_*\}\eta'(x) \times V^{-1}(x_1, \alpha_0)\eta(x)]$. Let $\varepsilon_n = n^{-1/2}b_n^{-s/4}$. The argument is similar as the proof of Lemma E.6. We write

$$\hat{D}_4 = \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \sum_{l=1, l\neq j\neq i}^{n} \mathbf{1}_{in} e_{ij} \rho(z_j, \alpha_0^n) \hat{V}(x_i, \hat{\alpha}_n)^{-1} \rho(z_l, \alpha_0^n) e_{il} + o_p(1)$$

under H_{1n}. Let $d_n(z_i, \alpha) = \rho(z_i, \alpha) - \varepsilon_n \eta(x_i)$, $i = 1, \ldots, n$. Then

$$\begin{aligned}
\hat{D}_4 \;&=\; \frac{1}{n^2 b_n^{2s}} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \sum_{l=1, l\neq j\neq i}^{n} \mathbf{1}_{in} K_{ij} d_n(z_j, \alpha_0^n)' \hat{H}(x_i, \hat{\alpha}_n)^{-1} d_n(z_l, \alpha_0^n) K_{il} \\
&\quad + \frac{\varepsilon_n^2}{n^2 b_n^{2s}} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \sum_{l=1, l\neq j\neq i}^{n} \mathbf{1}_{in} K_{ij} \eta(x_i)' \hat{H}(x_i, \hat{\alpha}_n)^{-1} \eta'(x_i) K_{il} + o_p(1) \\
&\equiv\; \hat{D}_{4a} + \hat{D}_{4b} + o_p(1).
\end{aligned}$$

First, note that $b_n^{s/2}\hat{D}_{4a} = o_p(1)$ by the consistency of kernel estimator and the consistency of

$\hat{\alpha}_n$. Second,

$$
\begin{aligned}
b_n^{s/2}\hat{D}_{4b} &= \frac{\varepsilon_n^2}{n^2 b_n^{2s}} \sum_{i=1}^n \sum_{j=1,j\neq i}^n \sum_{l=1,l\neq j\neq i}^n \mathbf{1}_{in} K_{ij}\eta(x_i)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\eta(x_i)K_{il} \\
&= \frac{1}{n^2 b_n^s} \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} K_{ij}\eta(x_i)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\left\{\eta(x_i)f(x_i)+r_{1i}\right\}
\end{aligned}
$$

where $r_{1i}=o_p(1)$ and the second equality follows from the definition of ε_n and the consistency of kernel estimators and (E.11). Hence,

$$
\begin{aligned}
b_n^{s/2}\hat{D}_{4b} &= \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in}\eta(x_i)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\eta(x_i)f^2(x_i) + \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in} r_{2i}'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\eta(x_i)f(x_i) \\
&\quad + \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in}\left\{\frac{1}{nb_n^s}\sum_{j=1,j\neq i}^n K_{ij}\eta(x_j)'\right\}\hat{H}(x_i,\hat{\alpha}_n)^{-1} r_{1i} \\
&= \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in}\eta(x_i)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\eta(x_i)f^2(x_i) + o_p(1) \\
&= \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in}\eta(x_i)'V(x_i,\alpha_0^n)^{-1}\eta(x_i) \\
&= \mu + o_p(1),
\end{aligned}
$$

where the second equality follows from the facts that r_{1i} and r_{2i} are asymptotically negligible element by element, the third equality follows from the consistency of $\hat{\alpha}_n$ to α_0^n and the last equality follows from the weak law of large numbers. Therefore, $b_n^{s/2}\hat{D}_4 \xrightarrow{d} N(\mu, 2qK^{**}vol(\mathcal{X}_n))$ and the desired result follows.

D Bootstrap Consistency

One may also be interested in a bootstrap version of SCELR_n, which is defined as

$$
\text{SCELR1}^\star \equiv -2\left\{\mathcal{L}_n^\star\left(\bar{\alpha}_n^\star\right) - \mathcal{L}_n^\star\left(\hat{\alpha}_n^\star\right)\right\},
$$

where $\bar{\alpha}_n^\star \equiv \left(\bar{\theta}_n^\star, \bar{h}_n^\star\right) \equiv \arg\max_{\alpha \in \mathcal{A}_n^{\hat{r}}} \mathcal{L}_n^\star(\alpha)$, $\hat{\alpha}_n^\star \equiv \left(\hat{\theta}_n^\star, \hat{h}_n^\star\right) \equiv \arg\max_{\alpha \in \mathcal{A}_n} \mathcal{L}_n^\star(\alpha)$, and $\mathcal{A}_n^{\hat{r}} \equiv \{\alpha \in \mathcal{A}_n : \mathcal{R}(\alpha) = \hat{r}\}$. A weighted bootstrap scheme draws bootstrap weights from an i.i.d sample of positive weights that satisfy $E[W_i] = 1$, $Var(W_i) = w_0$ and are independent of $\{(Y_i', X_i'\}_{i=1}^n$. We can obtain the convergence rate of the bootstrap estimators $\bar{\alpha}_n^\star$ and $\hat{\alpha}_n^\star$, respectively, and show that SCELR1* converges to chi-square distribution as well. To implement the bootstrap procedure, the confidence intervals are defined as

$$\mathcal{C}_n = \{\alpha \in \mathcal{A} : \text{SCELR1} \leq \hat{c}_{n,1-\tau}\},$$

where $\hat{c}_{n,1-\tau}$ is the $(1-\tau)$ quantile using the weighted bootstrap with variance 1. To simulate the critical value, for $b = 1, \ldots, B_n$, let

$$\hat{c}_{n,1-\tau} = \inf \left\{ t : \frac{1}{B_n} \sum_{b=1}^{B_n} \mathbf{1} \{\text{SCELR1}_b^\star \leq t\} \geq 1 - \tau \right\},$$

where $\text{SCELR1}_b^\star(\hat{r})$ is the bth bootstrap likelihood ratio statistic SCELR1*.

We highlight the consistency of bootstrap procedure as follows. Without loss of generality, we assume that $\mathcal{R}_l(\alpha)$ are linearly independent, i.e., $L = b$. Otherwise, we can always conduct a linear transformation for the hypothesis by the Gram-Schmidt orthogonalization. Let

$$l_n^\star(\alpha) = -\frac{1}{n} \sum_{i=1}^n W_i \mathbf{1}_{in} \sum_{j=1}^n e_{ij} \log(1 + \hat{\lambda}(x_i, \alpha)' \rho(z_j, \alpha)).$$

Define $\ddot{\alpha}^\star = \hat{\alpha}_n^\star - \sum_{l=1}^L \langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl} \rangle \tilde{v}_{nl} / \|\tilde{v}_{nl}\|^2$ and $\Psi_n \ddot{\alpha}^\star = \hat{\alpha}_n^\star - \sum_{l=1}^L \langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl} \rangle \Psi_n \tilde{v}_{nl} / \|\tilde{v}_{nl}\|^2$, where $\hat{\alpha}_n^\star$ is the maximizer of $l_n^\star(\alpha)$ over $\mathcal{A}_n$. The following steps are similar to the proof of Theorem 2.1. Note that the chain rule and Taylor expansion yield

$$l_n^\star(\hat{\alpha}_n^\star) - l_n^\star(\Psi_n \ddot{\alpha}^\star) = \frac{dl_n^\star(\Psi_n \ddot{\alpha}^\star)}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n \ddot{\alpha}^\star] + \frac{1}{2} \frac{d^2 l_n^\star(\tilde{\alpha}_n^\star)}{d\alpha d\alpha}[\hat{\alpha}_n^\star - \Psi_n \ddot{\alpha}^\star, \hat{\alpha}_n^\star - \Psi_n \ddot{\alpha}^\star], \tag{D.1}$$

where $\tilde{\alpha}_n^\star$ is a point in between $\hat{\alpha}_n^\star$ and $\Psi_n\ddot{\alpha}^\star$.

$$-\frac{dl_n^\star(\Psi_n\ddot{\alpha}^\star)}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star] = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{i,n} W_i \hat{\lambda}(x_i, \Psi_n\ddot{\alpha}^\star) \sum_{j=1}^n \frac{e_{ij}\frac{d\rho(z_j,\Psi_n\ddot{\alpha}^\star)}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star]}{1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha}^\star)'\rho(z_j,\Psi_n\ddot{\alpha}^\star)}$$

and

$$\begin{aligned}
&\frac{d^2 l_n^\star(\Psi_n\ddot{\alpha}^\star)}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star, \hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star] \\
= \;& \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{i,n} W_i \sum_{j=1}^n \frac{e_{ij}\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha}^\star)}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star]'\frac{d\rho(z_j,\Psi_n\ddot{\alpha}^\star)}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star]}{1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha}^\star)'\rho(z_j,\Psi_n\ddot{\alpha}^\star)}.
\end{aligned}$$

One can show that $\frac{dl_n^\star(\Psi_n\ddot{\alpha}^\star)}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star] = \sum_{l=1}^L \frac{\langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl}\rangle^2}{\|\tilde{v}_{nl}\|^2} + o_{p^\star}(n^{-1})$, and $\frac{d^2 l_n^\star(\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star, \hat{\alpha}_n^\star - \Psi_n\ddot{\alpha}^\star] = -\left(\sum_{l=1}^L \frac{\langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|}\right)^2 + o_p(1)$. Therefore,

$$l_n^\star(\hat{\alpha}_n^\star) - l_n^\star(\Psi_n\ddot{\alpha}^\star) = \frac{1}{2}\sum_{l=1}^L \left(\frac{\langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|}\right)^2 + o_{p^\star}(n^{-1}).$$

Next, define $\dddot{\alpha}^\star = \bar{\alpha}_n^\star + \sum_{l=1}^L \langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_l\rangle \tilde{v}_{nl} / \|\tilde{v}_{nl}\|^2$ and

$$\Psi_n\dddot{\alpha}^\star = \bar{\alpha}_n^\star + \sum_{l=1}^L \langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl}\rangle \Psi_n\tilde{v}_l / \|\tilde{v}_{nl}\|^2.$$

Following similar steps, we obtain

$$l_n^\star(\bar{\alpha}_n^\star) - l_n^\star(\Psi_n\dddot{\alpha}^\star) = -\frac{1}{2}\sum_{l=1}^L \left(\frac{\langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|}\right)^2 + o_{p^\star}(n^{-1}).$$

Since $\hat{\alpha}_n$ is a maximizer,

$$\begin{aligned}
l_n^\star(\hat{\alpha}_n^\star) - l_n^\star(\bar{\alpha}_n^\star) \;&\geq\; l_n^\star(\Psi_n\dddot{\alpha}^\star) - l_n^\star(\bar{\alpha}_n^\star) = \frac{1}{2}\sum_{l=1}^L \left(\frac{\langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|}\right)^2 + o_{p^\star}(n^{-1}) \\
&=\; l_n^\star(\hat{\alpha}_n^\star) - l_n^\star(\Psi_n\ddot{\alpha}^\star) + o_{p^\star}(n^{-1})
\end{aligned}$$

and therefore

$$l_n^\star(\Psi_n\ddot{\alpha}^\star) \geq l_n^\star(\bar{\alpha}_n^\star) + o_{p^\star}(n^{-1}).$$

Furthermore,

$$\begin{aligned} 2(\mathcal{L}_n^\star(\hat{\alpha}_n^\star) - \mathcal{L}_n^\star(\bar{\alpha}_n^\star)) &\leq 2(\mathcal{L}_n(\hat{\alpha}_n^\star) - \mathcal{L}_n(\Psi_n\ddot{\alpha}^\star(t))) + o_p(1) \\ &= 2n(l_n(\hat{\alpha}_n^\star) - l_n(\Psi_n\ddot{\alpha}^\star(t))) + o_p(1) \\ &= w_0 \sum_{l=1}^{L} \left(\frac{\sqrt{n}\,\langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl} \rangle}{\sqrt{w_0}\,\|\tilde{v}_{nl}\|} \right)^2 + o_{p^\star}(1) \end{aligned}$$

and

$$\begin{aligned} 2(\mathcal{L}_n^\star(\hat{\alpha}_n^\star) - \mathcal{L}_n^\star(\bar{\alpha}_n^\star)) &\geq 2(\mathcal{L}_n(\Psi_n\ddot{\alpha}^\star(t)) - \mathcal{L}_n(\bar{\alpha}_n^\star)) + o_p(1) \\ &= w_0 \sum_{l=1}^{L} \left(\frac{\sqrt{n}\,\langle \hat{\alpha}_n^\star - \hat{\alpha}_n, \tilde{v}_{nl} \rangle}{\sqrt{w_0}\,\|\tilde{v}_{nl}\|} \right)^2 + o_p(1) \end{aligned}$$

Since $\sqrt{n}\,\langle \hat{\alpha}_n - \alpha_0, \tilde{v}_{nl}/||\tilde{v}_{nl}|| \rangle \xrightarrow{d} N(0,1)$, $\sqrt{n}\,\left\langle \hat{\alpha}_n^* - \hat{\alpha}_n, \tilde{v}_{nl}/\sqrt{w_0}\,\|\tilde{v}_{nl}\| \right\rangle \xrightarrow{d} N(0,1)$, where $w_0 = Var(W-1) = Var(W)$.

$$2(\mathcal{L}_n^\star(\hat{\alpha}_n^*) - \mathcal{L}_n^\star(\bar{\alpha}_n^\star)) \xrightarrow{d} w_0\chi^2_{(b)}$$

where b is the maximum number of linearly independent constraints. □

E Useful Lemmas

In this section, we present some lemmas that are useful to prove the results in Section B.

Lemma E.1. *Suppose Assumptions in Theorem 2.1 hold, then*

(i)

$$\frac{dl_n(\Psi_n\ddot{\alpha})}{d\alpha}\left[\hat{\alpha}_n - \Psi_n\ddot{\alpha}\right] = \sum_{l=1}^{L} \frac{\langle \hat{\alpha}_n - \alpha_0, \tilde{v}_{nl} \rangle^2}{\|\tilde{v}_{nl}\|^2} + o_p(n^{-1}),$$

(ii)

$$\frac{d^2 l_n(\Psi_n\ddot{\alpha})}{d\alpha d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}, \hat{\alpha}_n - \Psi\ddot{\alpha}] = -\frac{1}{2}\sum_{l=1}^{L}\frac{\langle\hat{\alpha}_n - \alpha_0, \tilde{v}_{nl}\rangle^2}{\|\tilde{v}_{nl}\|^2} + o_p(n^{-1}).$$

Proof of E.1. (i) By the first order condition of (2.5), for all $\Psi_n\ddot{\alpha} \in \mathcal{N}_{0n}$ w.p.a.1, we have

$$0 = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ij}\rho(z_j, \Psi_n\ddot{\alpha})}{1 + \hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\rho(z_j, \Psi_n\ddot{\alpha})}, \tag{E.1}$$

A straightforward calculation shows that

$$\begin{aligned}
&\frac{dl_n(\Psi\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi\ddot{\alpha}] \\
= \quad & -\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{i,n}\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\sum_{j=1}^{n}\frac{e_{ij}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]}{1 + \hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\rho(z_j, \Psi_n\ddot{\alpha})} \\
= \quad & -\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\frac{d\hat{m}(x_i, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] \\
& + \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{i,n}\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\left(\sum_{j=1}^{n}\frac{e_{ij}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]\rho(z_j, \Psi_n\ddot{\alpha})'}{1 + \hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\rho(z_j, \Psi_n\ddot{\alpha})}\right)\hat{\lambda}(x_i, \Psi_n\ddot{\alpha}),
\end{aligned} \tag{E.2}$$

where the first equality follows from the first order condition and chain rule and the second equality follows from a straightforward calculation. Furthermore, by the definition of $\Psi_n\ddot{\alpha}$, $\Psi_n\ddot{\alpha} \in \mathcal{N}_{0n}$. Thus, respectively, Lemma B.4 (ii) and (iii) in Otsu (2011) implies that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\frac{d\hat{m}(x, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] \\
=&\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\hat{m}(x_i, \Psi_n\ddot{\alpha})'\hat{V}(x_i, \Psi_n\ddot{\alpha})^{-1}\frac{d\hat{m}(x_i, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] + o_p(n^{-1})
\end{aligned}$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{i,n}\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\left(\sum_{j=1}^{n}\frac{e_{ij}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]\rho(z_j, \Psi_n\ddot{\alpha})'}{1 + \hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\rho(z_j, \Psi_n\ddot{\alpha})}\right)\hat{\lambda}(x_i, \Psi_n\ddot{\alpha}) = o_p(n^{-1}).$$

Hence, (E.2) can be simplified as

$$
\begin{aligned}
& \frac{dl_n(\Psi\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi\ddot{\alpha}] \\
= & -\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{in}\hat{m}(x_i, \Psi_n\ddot{\alpha})'\hat{V}(x_i, \Psi_n\ddot{\alpha})^{-1}\frac{d\hat{m}(x_i, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] + o_p(n^{-1}) \\
= & -\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{in}\hat{m}(x_i, \Psi_n\ddot{\alpha})'V(x_i, \alpha_0)^{-1}\frac{dm(x_i, \alpha_0)}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] + o_p(n^{-1}) \\
= & -\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{in}\hat{m}(x_i, \alpha_0)'V(x_i, \alpha_0)^{-1}\frac{dm(x_i, \alpha_0)}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] \\
& -\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{in}\left\{\hat{m}(x_i, \Psi_n\ddot{\alpha}) - \hat{m}(x_i, \alpha_0)\right\}' V(x_i, \alpha_0)^{-1}\frac{dm(x_i, \alpha_0)}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] + o_p(n^{-1}) \\
= & -\sum_{l=1}^{L} \frac{\langle \hat{\alpha}_n - \alpha_0, \Psi_n\tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|^2} \times \Bigg(\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{in}\rho(z_i, \alpha_0)'V(x_i, \alpha_0)^{-1}\frac{dm(x_i, \alpha_0)}{d\alpha}[\tilde{v}_{nl}] \\
& + \langle \Psi_n\ddot{\alpha} - \alpha_0, \tilde{v}_{nl}\rangle + o_p(n^{-1})\Bigg) \\
= & -\sum_{l=1}^{L} \frac{\langle \hat{\alpha}_n - \alpha_0, \tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|^2} \times \left(-\langle \hat{\alpha}_n - \alpha_0, \tilde{v}_{nl}\rangle + o_p(n^{-1})\right) \\
= & \sum_{l=1}^{L} \frac{\langle \hat{\alpha}_n - \alpha_0, \tilde{v}_{nl}\rangle^2}{\|\tilde{v}_{nl}\|^2} + o_p(n^{-1}),
\end{aligned}
$$

where the second equality follows from Lemma (iv) in Otsu (2011), forth equality follows from the definition of $\Psi_n\ddot{\alpha}$ and Lemma B.4 (v) in Otsu (2011), and the fifth equality follows from the property that $\langle \Psi_n\tilde{v}_{nl} - \tilde{v}_{nl}, \tilde{v}_{nl}/\|\tilde{v}_{nl}\|\rangle = o_p(1)$ by assumptions in Theorem 2.1.

(ii) Next, we decompose the second-order term as

$$
\frac{d^2 l_n(\Psi_n\ddot{\alpha})}{d\alpha d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}, \hat{\alpha}_n - \Psi_n\ddot{\alpha}] = A_1 + A_2 + A_3,
$$

where

$$
\begin{aligned}
A_1 &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\mathbf{1}_{in}e_{ij}\left\{\frac{d\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\rho(z_j, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]\right\}\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\frac{d\rho(z_j, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]}{\left(1 + \hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\rho(z_j, \Psi_n\ddot{\alpha})\right)^2}, \\
A_2 &= -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\mathbf{1}_{in}e_{ij}\frac{d\hat{\lambda}(x_i, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]'\frac{d\rho(z_j, \Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]}{1 + \hat{\lambda}(x_i, \Psi_n\ddot{\alpha})'\rho(z_j, \Psi_n\ddot{\alpha})},
\end{aligned}
$$

and

$$A_3 = - \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\mathbf{1}_{in}e_{ij}}{1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})} \times \sum_{k=1}^{d_\rho}\left\{\frac{d^2\rho^{(k)}(z_j,\Psi_n\ddot{\alpha})}{d\alpha d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha},\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right\}\hat{\lambda}^{(k)}(x_i,\Psi_n\ddot{\alpha}).$$

The desired result follows by the following steps.

First, we show that $\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}|A_1(\Psi_n\ddot{\alpha})|$ is $o_p(n^{-1})$. We write

$$\begin{aligned}
&\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}|A_1(\Psi_n\ddot{\alpha})| \\
= &\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\left|\sum_{l=1}^{L}\frac{\langle\hat{\alpha}_n-\alpha_0,\Psi_n\tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|}\right. \\
&\left.\times\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\mathbf{1}_{in}e_{ij}\left\{\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\tilde{v}_{nl}]\right\}\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\tilde{v}_{nl}]}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)^2}\right)\right|,
\end{aligned}$$

where

$$\begin{aligned}
&\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\max_{1\le l\le L}\left|\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\mathbf{1}_{in}e_{ij}\left\{\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\tilde{v}_{nl}]\right\}\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\tilde{v}_{nl}]}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)^2}\right| \\
\le &\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\max_{1\le l\le L}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}e_{ij}\left\{\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\tilde{v}_{nl}]+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\tilde{v}_{nl}]\right\}\right. \\
&\left.\times\ \hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\tilde{v}_{nl}]\right|\times O_{as}(1)
\end{aligned}$$

where the inequality follows because $\max_{1\le j\le n}\sup_{(x_i,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\left|\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)\right| = o_{as}(1)$. Thus, we have

$$
\begin{aligned}
& \sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}} |A_1(\Psi_n\ddot{\alpha})| \\
\leq\ & O_{as}(1)\times \sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}} \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\times\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}\left[\hat{\alpha}_n-\Psi_n\ddot{\alpha}\right]\right\|_E \\
& \times \sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}}\left\|\sum_{j=1}^{n}e_{ij}\rho(z_j,\Psi_n\ddot{\alpha})'\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right\|_E \\
& +O_{as}(1)\times \sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})\right\|_E^2 \\
& \times \sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}}\left\|\sum_{j=1}^{n}e_{ij}\left\{\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right\}'\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right\|_E .
\end{aligned}
$$

It implies that

$$
\begin{aligned}
& \sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}} |A_1(\Psi_n\ddot{\alpha})| \\
\leq\ & O_{as}(1)\times \sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}} \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\sum_{l=1}^{L}\frac{\langle\hat{\alpha}_n-\alpha_0,\Psi_n\tilde{v}_l\rangle}{\|\tilde{v}_{nl}\|^2}\hat{\lambda}(x_i,\hat{\alpha}_n)'\times\frac{d\hat{\lambda}(x_i,\hat{\alpha}_n)}{d\alpha}[\tilde{v}_{nl}]\right\|_E \\
& \times \sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}}\left\|\sum_{l=1}^{L}\left(\frac{\langle\hat{\alpha}_n-\alpha_0,\Psi_n\tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|}\sum_{j=1}^{n}e_{ji}\rho(z_j,\hat{\alpha}_n)'\frac{d\rho(z_j,\hat{\alpha}_n)}{d\alpha}[\tilde{v}_{nl}]\right)\right\|_E \\
& +O_{as}(1)\times \sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\hat{\lambda}(x_i,\hat{\alpha}_n)\right\|_E^2 \\
& \times \sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}}\left\|\sum_{l=1}^{L}\left(\frac{\langle\hat{\alpha}_n-\alpha_0,\Psi_n\tilde{v}_l\rangle^2}{\|\tilde{v}_{nl}\|^2}\sum_{j=1}^{n}e_{ij}\left\{\frac{d\rho(z_j,\hat{\alpha}_n)}{d\alpha}[\tilde{v}_{nl}]\right\}'\frac{d\rho(z_j,\hat{\alpha}_n)}{d\alpha}[\tilde{v}_{nl}]\right)\right\|_E \\
\leq\ & O_p(n^{-1})\times o_p(n^{-1/4})\sup_{x_i\in\mathcal{X}_n}\|\sum_{j=1}^{n}e_{ij}c_1(z_j)c_2(z_j)\|_E+o_p(n^{-1/2})\times O_p(n^{-1})\sup_{x_i\in\mathcal{X}_n}\|\sum_{j=1}^{n}e_{ij}c_2(z_j)^2\|_E \\
=\ & o_p(n^{-1}),
\end{aligned}
$$

where the second inequality follows from the fact that for each $l=1,\ldots,L$, $\langle\hat{\alpha}_n-\alpha_0,\Psi_n\tilde{v}_{nl}\rangle/\|\tilde{v}_{nl}\|=$

$O_p(n^{-1/2})$, $\sup_{\tilde{\alpha}_n \in \mathcal{N}_{0n}} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{in} \left\| \hat{\lambda}(x_i, \tilde{\alpha}_n) \right\|_E^2 = o_p(n^{-1/2})$,

$$\sup_{\tilde{\alpha}_n \in \mathcal{N}_{0n}} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{in} \left\| \hat{\lambda}(x_i, \tilde{\alpha}_n)' \times \frac{d\hat{\lambda}(x_i, \tilde{\alpha}_n)}{d\alpha}[\tilde{v}_{nl}] \right\|_E = o_p(n^{-1/4})$$

and Assumptions A.2 and A.8, where $c_1(z_j)$ and $c_2(z_j)$ are envelope functions for $\rho(z_j, \tilde{\alpha}_n)$ and $\frac{d\rho(z_j,\tilde{\alpha}_n)}{d\alpha}[\tilde{v}_{nl}]$, respectively. And the last equality follows from the uniform convergence of kernel estimators.

Next, we show that $\sup_{\Psi_n \ddot{\alpha} \in \mathcal{N}_{0n}} |A_3(\Psi_n \ddot{\alpha})| = o_p(n^{-1})$. Similarly, we have

$$\begin{aligned}
& \sup_{\hat{\alpha}_n \in \mathcal{N}_{0n}} |A_3(\Psi_n \ddot{\alpha})| \\
\leq\; & \max_{1 \leq j \leq n} \sup_{(x_i, \Psi_n \ddot{\alpha}) \in \mathcal{X}_n \times \mathcal{N}_{0n}} \left| \frac{1}{1 + \hat{\lambda}(x_i, \Psi_n \ddot{\alpha})' \rho(z_j, \Psi_n \ddot{\alpha})} \right| \times \sup_{\Psi_n \ddot{\alpha} \in \mathcal{N}_{0n}} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{in} \left\| \max_{1 \leq k \leq d_\rho} \lambda_k(x_i, \hat{\alpha}_n) \right\|_E \\
& \sup_{(x_i, \Psi_n \ddot{\alpha}) \in \mathcal{X}_n \times \mathcal{N}_{0n}} \left\| \sum_{l=1}^L \left\{ \left(\frac{\langle \hat{\alpha}_n - \alpha_0, \Psi_n \tilde{v}_l \rangle}{\|\tilde{v}_{nl}\|} \right)^2 \sum_{j=1}^n e_{ij} \sum_{k=1}^{d_\rho} \frac{d^2 \rho_k(z_j, \Psi_n \ddot{\alpha})}{d\alpha d\alpha}[\tilde{v}_{nl}, \tilde{v}_{nl}] \right\} \right\|_E \\
\leq\; & O_{as}(1) \times O_p(n^{-1}) \times O_p(n^{-1/2}) \times \sup_{\hat{\alpha}_n \in \mathcal{N}_{0n}} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{in} \left\| \max_{1 \leq k \leq d_\rho} \lambda_k(x_i, \hat{\alpha}_n) \right\|_E \times \sup_{x_i \in \mathcal{X}_n} \left\| \sum_{j=1}^n e_{ji} c_3(z_j) \right\|_E \\
=\; & o_p(n^{-1})
\end{aligned}$$

where the first inequality is by direct calculation, the second inequality follows from the fact that for each $l = 1, \ldots, L$, $\langle \hat{\alpha}_n - \alpha_0, \Psi_n \tilde{v}_{nl} \rangle / \|\tilde{v}_{nl}\| = O_p(n^{-1/2})$,$\max_{1 \leq j \leq n} \sup_{(x_i, \alpha) \times \mathcal{X}} \left| \hat{\lambda}(x_i, \alpha)' \rho(z_j, \alpha) \right| = o_{as}(1)$ by Lemma D.2 in Kitamura *et al.* (2004) and Assumption A.8, where $c_3(z_j)$ is an envelope function for $\frac{d^2 \rho_k(z_j, \Psi_n \ddot{\alpha})}{d\alpha d\alpha}[\tilde{v}_{nl}, \tilde{v}_{nl}]$, and the last equality follows from the uniform convergence of kernel estimators and the fact that $\sup_{\hat{\alpha}_n \in \mathcal{N}_{0n}} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{in} \|\lambda(x_i, \hat{\alpha}_n)\|_E = o_p(n^{-1/4})$.

Next, we prove that $A_2 = -\frac{1}{2n}\sum_{l=1}^{L}\left(\frac{\langle \hat{\alpha}_n - \alpha_0, \Psi_n \tilde{v}_l\rangle}{\|\tilde{v}_l\|}\right)^2 + o_p(1)$. Note that by equation (E.1)

$$
\begin{aligned}
0 &= \frac{d}{d\alpha}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ij}\rho(z_j,\Psi_n\ddot{\alpha})}{1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})}\right)[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ij}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)} \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ij}\rho(z_j,\Psi_n\ddot{\alpha})\rho(z_j,\Psi_n\ddot{\alpha})'}{\left(1+\hat{\lambda}(x_i,\hat{\alpha}_n)'\rho(z_j,\hat{\alpha}_n)\right)^2}\times\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}] \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ij}\rho(z_j,\Psi_n\ddot{\alpha})\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]'}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)^2}\times\hat{\lambda}(x_i,\Psi_n\ddot{\alpha}) \\
&\equiv B_1 - B_2 - B_3, \qquad\qquad (E.3)
\end{aligned}
$$

Similarly, some algebra shows that

$$
\begin{aligned}
B_3 &\le \sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left(\sum_{j=1}^{n}\frac{e_{ji}\rho(z_j,\Psi_n\ddot{\alpha})\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]'}{\left(1+\hat{\lambda}(x_i,\hat{\alpha}_n)'\rho(z_j,\hat{\alpha}_n)\right)^2}\right)\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})\right| \\
&\le \max_{1\le j\le n}\sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}}\left|\frac{1}{\left(1+\hat{\lambda}(x_i,\hat{\alpha}_n)'\rho(z_j,\hat{\alpha}_n)\right)^2}\right| \\
&\quad\times\sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}}\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})\right\|_E\times\left\|\sum_{j=1}^{n}e_{ij}\rho(z_j,\Psi_n\ddot{\alpha})\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}]\right\|_E \\
&\le O_p(n^{-1/2})\times O_{as}(1)\times\sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}}\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})\right\|_E\|\hat{m}(x_i,\Psi_n\ddot{\alpha})\|_E \\
&\quad\times O_p(1) = O_p(n^{-1/2})\times o_p(n^{-1/2}) = o_p(n^{-1}), \qquad\qquad (E.4)
\end{aligned}
$$

where the last equality follows because

$$\begin{aligned}
&\sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}} \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})\right\|_E \|\hat{m}(x_i,\Psi_n\ddot{\alpha})\|_E \\
\leq &\sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}} \sqrt{\frac{1}{n}\sum_{i=1}^{n}C^{-1}\|\hat{m}(x_i,\Psi_n\ddot{\alpha})\|_E^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\|\hat{m}(x_i,\Psi_n\ddot{\alpha})\|_E^2} \\
\simeq &\sup_{(x_i,\Psi_n\ddot{\alpha})\in\mathcal{X}_n\times\mathcal{N}_{0n}} \frac{1}{n}\sum_{i=1}^{n}\|\hat{m}(x_i,\Psi_n\ddot{\alpha})\|_E^2 = o_p(n^{-1/4}),
\end{aligned}$$

where the first inequality follows from Cauchy-Schwartz and the fact that $\|\hat{\lambda}(x_i,\alpha)\|_E \leq C^{-1}\|\hat{m}(x_i,\alpha)\|_E$ for all $\alpha\in\mathcal{N}_{0n}$, and the second equality follows from similar argument in Lemma E.7. Next, we show that

$$B_2 = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}V(x_i,\Psi_n\ddot{\alpha})\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]+o_p(n^{-1}). \tag{E.5}$$

The argument is as follows.

$$\begin{aligned}
&\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\left(\frac{e_{ji}\rho(z_j,\Psi_n\ddot{\alpha})\rho(z_j,\Psi_n\ddot{\alpha})'}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)^2}\right)\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right| \\
&-\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}V(x_i,\Psi_n\ddot{\alpha})\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right| \\
\leq\;&\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\left\{\left(\frac{e_{ji}\rho(z_j,\Psi_n\ddot{\alpha})\rho(z_j,\Psi_n\ddot{\alpha})'}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)^2}\right)-V(x_i,\Psi_n\ddot{\alpha})\right\}\right. \\
&\left.\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right|\times O_p(n^{-1/2}) \\
\leq\;&O_p(n^{-1/2})\times O_p(1) \\
&\times\sup_{\Psi_n\ddot{\alpha}\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\left\|\hat{V}(x_i,\Psi_n\ddot{\alpha})-V(x_i,\Psi_n\ddot{\alpha})\right\|_E\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right| \\
=\;&O_p(n^{-1/2})\times O_p(1)\times o_p(n^{-1/4})\times o_p(n^{-1/4})=o_p(n^{-1}).
\end{aligned}$$

By combining results in equations (E.3), (E.4) and (E.5), we have

$$\begin{aligned}
B_1 &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ji}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)} = B_2 + o_p(n^{-1/2}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}V(x_i,\Psi_n\ddot{\alpha})\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}] + o_p(n^{-1/2}).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}] \\
&= V(x_i,\Psi_n\ddot{\alpha})^{-1}\sum_{j=1}^{n}\left\{\frac{e_{ji}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)}\right\} + o_p(n^{-1}).
\end{aligned}$$

Thus, we obtain the result that

$$\begin{aligned}
&A_2 \\
&= -\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\frac{d\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]'\sum_{j=1}^{n}\frac{\mathbf{1}_{i,n}e_{ij}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]}{1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})} \\
&= -\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\left\{\frac{e_{ji}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]}{\left(1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})\right)}\right\}V(x_i,\Psi_n\ddot{\alpha})^{-1} \\
&\quad \sum_{j=1}^{n}\left\{\frac{\mathbf{1}_{i,n}e_{ij}\frac{d\rho(z_j,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]}{1+\hat{\lambda}(x_i,\Psi_n\ddot{\alpha})'\rho(z_j,\Psi_n\ddot{\alpha})}\right\} + o_p(n^{-1}) \\
&= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{d\hat{m}(z_i,\hat{\alpha}_n)}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}]\right)'V(x_i,\Psi_n\ddot{\alpha})^{-1}\frac{d\hat{m}(z_i,\Psi_n\ddot{\alpha})}{d\alpha}[\hat{\alpha}_n-\Psi_n\ddot{\alpha}] + o_p(n^{-1}) \\
&= -\sum_{l=1}^{L}\left\{\left(\frac{\langle\hat{\alpha}_n-\alpha_0,\tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|^2}\right)^2\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{dm(z_i,\alpha_0)}{d\alpha}[\tilde{v}_{nl}]\right)'V(x_i,\alpha_0)^{-1}\frac{dm(z_i,\alpha_0)}{d\alpha}[\tilde{v}_{nl}]\right]\right\} + o_p(n^{-1}) \\
&= -\sum_{l=1}^{L}\left(\frac{\langle\hat{\alpha}_n-\alpha_0,\tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|}\right)^2\times\left\{\|\tilde{v}_{nl}\|^2+o_P(n^{-1/2})\right\} + o_p(n^{-1}) \\
&= -\sum_{l=1}^{L}\left(\frac{\langle\hat{\alpha}_n-\alpha_0,\tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|^2}\right)^2 + o_p(n^{-1}),
\end{aligned}$$

where the third equation follows from the fact that $\max_{1\le j\le n}\sup_{(x_i,\alpha)\times\mathcal{X}}\left|\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)\right| =$

$o_{as}(1)$, the forth equation follows from Lemma B.2 (ii) and (iii) in Otsu (2011), and the fifth equation follows from the definition of $\hat{\Psi}_n\ddot{\alpha}$. Hence, the result

$$\frac{1}{2}\frac{d^2 l_n(\Psi_n\ddot{\alpha})}{d\alpha d\alpha}[\hat{\alpha}_n - \Psi_n\ddot{\alpha}, \hat{\alpha}_n - \Psi\ddot{\alpha}] = -\frac{1}{2}\sum_{l=1}^{L}\frac{\langle\hat{\alpha}_n - \alpha_0, \tilde{v}_{nl}\rangle}{\|\tilde{v}_{nl}\|^2} + o_p(n^{-1}) \tag{E.6}$$

follows. □

Lemma E.2. *Suppose that Assumptions A.1-A.9 hold. Then*

$$\sup_{\alpha\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ij}\rho(z_j,\alpha)\left(\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)\right)^2}{1+\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)}\right| = o_p(n^{-1/2}).$$

Proof.

$$\begin{aligned}
&\sup_{\alpha\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\sum_{j=1}^{n}\frac{e_{ij}\rho(z_j,\alpha)\left(\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)\right)^2}{1+\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)}\right| \\
\leq &\max_{1\leq j\leq n}\sup_{(x_i,\alpha)\in\mathcal{X}_n\times\mathcal{N}_{0n}}\left|\frac{1}{1+\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)}\right|\sup_{\alpha\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\|\hat{\lambda}(x_i,\alpha)\|_E^2\right| \\
&\times\sup_{(x_i,\alpha)\in\mathcal{X}_n\times\mathcal{N}_{0n}}\left|\sum_{j=1}^{n}e_{ij}\|\rho(z_j,\alpha)\|^3\right| \\
\leq &O_p(1)\sup_{x_i\in\mathcal{X}_n}\left|\sum_{j=1}^{n}e_{ij}c_1(z_j)^3\right|\sup_{\alpha\in\mathcal{N}_{0n}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{in}\|\hat{\lambda}(x_i,\alpha)\|_E^2\right| = o_p(n^{-1/2}),
\end{aligned}$$

where the second inequality follows from the fact that $\max_{1\leq j\leq n}\sup_{(x_i,\alpha)}\left|\hat{\lambda}(x_i,\alpha)'\rho(z_j,\alpha)\right| = o_{as}(1)$, Assumption A.2 ($c_1(z_j)$ is an envelope function for $\rho(z_j,\alpha)$) and the last equality follows from Lemma B.3 (vii) in Otsu (2011), the uniform convergence of kernel estimators and Lemma D.4 of Kitamura *et al.* (2004). □

Lemma E.3. *Suppose that Assumptions A.1-A.9 hold. Then $b_n^{s/2}\hat{D}_1 = o_p(1)$ under H_0.*

Proof. Recall that $\hat{H}(x_i,\hat{\alpha}_n) = \hat{V}(x_i,\hat{\alpha}_n)\hat{f}^2(x_i)$ and $\hat{f}(x_i) = \sum_{u=1}^{n}K_{iu}/nb_n^s$. By the uniform consistency of kernel estimators, it is well-known that

$$\sup_{x\in\mathcal{X}_n}|\hat{f}^2(x)-f^2(x)|=O_p\left(\sqrt{\frac{\log n}{nb_n^s}}+b_n^2\right). \tag{E.7}$$

Also, by Lemma B.2 in Otsu (2011), we have

$$\sup_{(x,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\|\hat{V}(x_i,\alpha)-V(x_i,\alpha)\|_E=O_p\left(\varrho_n\right)=o_p(n^{-1/4}), \tag{E.8}$$

Let $\inf_{x\in\mathcal{X}_n}|f(x)|=b$ and $\sup_{(x,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\|V(x_i,\alpha)\|=B$, for constants b and B satisfying $0<b,B<\infty$. By some algebra,

$$\begin{aligned}
&\sup_{(x,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\left\|\hat{H}(x,\alpha)-H(x,\alpha)\right\| \\
=&\sup_{(x,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\left\|\frac{\hat{V}(x,\alpha)}{\hat{f}^2(x)}-\frac{V(x,\alpha)}{f^2(x)}\right\| &&\text{(E.9)}\\
\leq& b^{-1}\sup_{(x,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\left\|\hat{V}(x,\alpha)-V(x,\alpha)\right\|+\frac{B(1+2/b)}{b^2}\sup_{x\in\mathcal{X}_n}\left|\hat{f}^2(x)-f^2(x)\right| \\
&+(1+2/b)/b^2\sup_{(x,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\left\|\hat{V}(x,\alpha)-V(x,\alpha)\right\|\times\sup_{x\in\mathcal{X}_n}\left|\hat{f}^2(x)-f^2(x)\right| &&\text{(E.10)}
\end{aligned}$$

Thus, (E.7)-(E.10) imply that

$$\sup_{(x,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}\left\|\hat{H}(x,\alpha)-H(x,\alpha)\right\|=O_p\left(\sqrt{\frac{\log n}{nb_n^s}}+b_n^2+\varrho_n\right). \tag{E.11}$$

Then

$$\begin{aligned}
\sup_{(x_i,\hat{\alpha}_n)\in\mathcal{X}_n\times\mathcal{N}_{0n}}b_n^{s/2}\left|\hat{D}_1\right| &= \sup_{(x_i,\hat{\alpha}_n)\in\mathcal{X}_n\times\mathcal{N}_{0n}}b_n^{s/2}\left|K^2(0)\sum_{i=1}^n\mathbf{1}_{in}\frac{\rho(z_i,\hat{\alpha}_n)'\hat{V}(x_i,\hat{\alpha}_n)^{-1}\rho(z_i,\hat{\alpha}_n)}{\{\sum_{u=1}^n K_{iu}\}^2}\right| \\
&= \sup_{(x_i,\hat{\alpha}_n)\in\mathcal{X}_n\times\mathcal{N}_{0n}}b_n^{s/2}\left|\frac{K^2(0)}{n^2b_n^{2s}}\sum_{i=1}^n\mathbf{1}_{in}\rho(z_i,\hat{\alpha}_n)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\rho(z_i,\hat{\alpha}_n)\right| \\
&\leq K^2(0)\sup_{(x_i,\hat{\alpha}_n)\in\mathcal{X}_n\times\mathcal{N}_{0n}}b_n^{s/2}\left\|\hat{H}(x_i,\hat{\alpha}_n)^{-1}\right\|\frac{1}{n^2b_n^{2s}}\sum_{i=1}^n c_1(z_i)^2 \\
&= O_p(n^{-1}b_n^{-3s/2}).
\end{aligned}$$

We have $b_n^{s/2}\hat{D}_1 = o_p(1)$ by Assumption A.3. $\square$

Lemma E.4. *Suppose that Assumptions A.1-A.9 hold. Then* $b_n^{s/2}\hat{D}_2 = b_n^{-s/2} qR(K)vol(\mathcal{X}_n) + o_p(1)$ *under* H_0.

Proof. Recall that $e_{ij} = K_{ij}/\sum_{j=1}^n K_{ij}$. We decompose

$$\hat{D}_2 = \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} e_{ij}^2 \rho(z_j,\hat{\alpha}_n)'\hat{V}(x_i,\hat{\alpha}_n)^{-1}\rho(z_j,\hat{\alpha}_n)$$

as

$$\begin{aligned}
\hat{D}_2 \;=\; & \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} e_{ij}^2 \rho(z_j,\alpha_0)'\hat{V}(x_i,\hat{\alpha}_n)^{-1}\rho(z_j,\alpha_0) \\
& + \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} e_{ij}^2 \left\{\rho(z_j,\hat{\alpha}_n) - \rho(z_j,\alpha_0)\right\}' \hat{V}(x_i,\hat{\alpha}_n)^{-1} \left\{\rho(z_j,\hat{\alpha}_n) - \rho(z_j,\alpha_0)\right\} \\
& + \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} e_{ij}^2 \rho(z_j,\alpha_0)'\hat{V}(x_i,\hat{\alpha}_n)^{-1} \left\{\rho(z_j,\hat{\alpha}_n) - \rho(z_j,\alpha_0)\right\} \\
& + \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} e_{ij}^2 \left\{\rho(z_j,\hat{\alpha}_n) - \rho(z_j,\alpha_0)\right\}' \hat{V}(x_i,\hat{\alpha}_n)^{-1}\rho(z_j,\alpha_0) \\
\equiv\; & \hat{D}_{2a} + \hat{D}_{2b} + \hat{D}_{2c} + \hat{D}_{2d}.
\end{aligned}$$

Let us look at each term of $\hat{D}_2$ one by one. First,

$$\begin{aligned}
\hat{D}_{2a} \;=\; & \frac{1}{n^2 b_n^{2s}} tr \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} K_{ij}^2 \rho(z_j,\alpha_0)' \left\{\hat{H}(x_i,\hat{\alpha}_n)^{-1} - H(x_i,\alpha_0)^{-1}\right\} \rho(z_j,\alpha_0) \\
& + \frac{1}{n^2 b_n^{2s}} tr \sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{1}_{in} K_{ij}^2 \rho(z_j,\alpha_0)' H(x_i,\alpha_0)^{-1}\rho(z_j,\alpha_0) \equiv \hat{D}_{2a}^{(1)} + \hat{D}_{2a}^{(2)}.
\end{aligned}$$

First, by similar argument in Lemma E.3, one can show that $b_n^{s/2}\hat{D}_{2a}^{(1)} = o_p(b_n^{-s/2} n^{-1/4}) = o_p(1)$.

Next, for the second term, note that

$$\begin{aligned}\hat{D}_{2a}^{(2)} =& \frac{1}{nb_n^s} tr \sum_{i=1}^{n} \mathbf{1}_{in} \left\{ R(K) V(x_i, \alpha_0) f(x_i) + R_a(x_i) \right\} H^{-1}(x_i, \alpha_0) \\ =& \frac{qR(K)}{nb_n^s} \sum_{i=1}^{n} \frac{\mathbf{1}_{in}}{f(x_i)} + O_p\left(\frac{\zeta_n}{b_n^s}\right),\end{aligned}$$

where q is the number of moments $\rho(x_i, \alpha_0)$, $R(K) = \int_{[-1,1]^s} K^2(u)du$, $\sup_{x \in \mathcal{X}_n} \|R_a(x_i)\| = O_p(\zeta_n)$, $\zeta_n = \sqrt{\frac{\log n}{nb_n^s}} + b_n^2$ follows by the uniform consistency of kernel estimators and the second equation holds because $\sup_{x \in \mathcal{X}_n} \|H^{-1}(x_i, \alpha_0)\| < \infty$. Furthermore, $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\mathbf{1}_{in}}{f(x_i)} = vol(\mathcal{X}_n) + o_p(1)$ by the central limit theorem, which implies that $b_n^{s/2} \hat{D}_{2a}^{(1)} = b_n^{-s/2} qR(K) vol(\mathcal{X}_n) + o_p(1)$.

Next, note that $\sup_{(x_i, \hat{\alpha}_n) \in \mathcal{X}_n \times \mathcal{N}_{0n}} \|\rho(z_j, \hat{\alpha}_n) - \rho(z_j, \alpha_0)\| = O_p(\varrho_{sn})$ by Assumption 2.1 and Assumption A.2, then

$$\hat{D}_{2b} \leq \frac{1}{nb_n^s} tr \sum_{i=1}^{n} \mathbf{1}_{in} \left\{ \frac{1}{nb_n^s} \sum_{j=1, j \neq i}^{n} K_{ij}^2 \varrho_{sn}^2 \right\} \hat{H}(x_i, \hat{\alpha}_n)^{-1} = O_p\left(\varrho_{sn}^2 b_n^{-s}\right).$$

Thus, $b_n^{s/2} \hat{D}_{2b} = o_p(1)$ by consistency of kernel estimator, consistency of $\hat{\alpha}_n$ and Assumption A.3. Similarly, we have $\hat{D}_{2c} = o_p(b_n^{-s/2})$ and $\hat{D}_{2d} = o_p(b_n^{-s/2})$, which implies, in summary, $b_n^{s/2} \hat{D}_2 = b_n^{-s/2} qR(K) vol(\mathcal{X}_n) + o_p(1)$. □

Lemma E.5. *Suppose Assumption A.1-A.9 hold. Then* $b_n^{s/2} \hat{D}_3 = o_P(1)$.

Proof. First, we rewrite $\hat{D}_3$ as the following terms:

$$\begin{aligned}
\frac{1}{2}\hat{D}_3 &= K(0)\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}\frac{\rho(z_i,\hat{\alpha}_n)'\hat{V}(x_i,\hat{\alpha}_n)^{-1}\rho(z_j,\hat{\alpha}_n)e_{ij}}{\sum_{u=1}^{n}K_{iu}} \\
&= \frac{K(0)}{n^2b_n^{2s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\rho(z_i,\hat{\alpha}_n)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\rho(z_j,\hat{\alpha}_n) \\
&= \frac{K(0)}{n^2b_n^{2s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\rho(z_i,\alpha_0)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\rho(z_j,\alpha_0) \\
&\quad +\frac{2K(0)}{n^2b_n^{2s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\rho(z_i,\alpha_0)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\left\{\rho(z_j,\hat{\alpha}_n)-\rho(z_j,\alpha_0)\right\} \\
&\quad +\frac{K(0)}{n^2b_n^{2s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\left\{\rho(z_i,\hat{\alpha}_n)-\rho(z_i,\alpha_0)\right\}'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\left\{\rho(z_j,\hat{\alpha}_n)-\rho(z_j,\alpha_0)\right\} \\
&\equiv \hat{D}_{3a}+\hat{D}_{3b}+\hat{D}_{3c}.
\end{aligned}$$

Let $\hat{\psi}^{(st)}(x_i)$ be the $(st)^{th}$ element of $\hat{H}(x_i,\hat{\alpha}_n)^{-1}$ and $\psi^{(st)}$ be the $(st)^{th}$ element of $H(x_i,\alpha_0)^{-1}$. We write the first term as

$$\begin{aligned}
&\hat{D}_{3a} \\
=\ & \frac{K(0)}{n^2b_n^{2s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\rho(z_i,\alpha_0)'\hat{H}(x_i,\hat{\alpha}_n)^{-1}\rho(z_j,\alpha_0) \\
=\ & \sum_{s=1}^{q}\sum_{t=1}^{q}\frac{K(0)}{n^2b_n^{2s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\rho^{(s)}(z_i,\alpha_0)\hat{\psi}^{(st)}(x_i,\hat{\alpha}_n)^{-1}\rho^{(t)}(z_j,\alpha_0) \\
=\ & \frac{K(0)}{n^{1/2}b_n^{s}}\sum_{s=1}^{q}\sum_{t=1}^{q}\left\{\frac{1}{n^{3/2}b_n^{s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\rho^{(s)}(z_i,\alpha_0)\left[\hat{\psi}^{(st)}(x_i,\hat{\alpha}_n)^{-1}-\psi^{(st)}(x_i,\alpha_0)^{-1}\right]\rho^{(t)}(z_j,\alpha_0)\right\} \\
& +\frac{K(0)}{n^{1/2}b_n^{s}}\sum_{s=1}^{q}\sum_{t=1}^{q}\left\{\frac{1}{n^{3/2}b_n^{s}}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbf{1}_{in}K_{ij}\rho^{(s)}(z_i,\alpha_0)\psi^{(st)}(x_i,\alpha_0)^{-1}\rho^{(t)}(z_j,\alpha_0)\right\} \\
\equiv\ & \frac{K(0)}{n^{1/2}b_n^{s}}\sum_{s=1}^{q}\sum_{t=1}^{q}\left(\hat{D}_{3a}^{(1)}+\hat{D}_{3a}^{(2)}\right).
\end{aligned}$$

Because $\sum_{i=1}^{n}\left(\frac{1}{nb_n^s}\sum_{j=1,j\neq i}^{n}\rho^{(t)}(z_j,\alpha_0)K_{ij}\right)^2 = O(b_n^{-s})$ by the fact that $K_{ij}\rho^{(s)}(z_i,\alpha_0)$ and

$K_{ik}\rho^{(t)}(Z_j, \alpha_0)$ are uncorrelated for $i \neq j \neq k$ and

$$\sup_{(x_i, \alpha_n) \in \mathcal{X}_n \times \mathcal{N}_{0n}} \left| \hat{\psi}^{(st)}(x_i, \hat{\alpha}_n) - \psi^{(st)}(x_i, \alpha_0) \right|^2 = O_p\left(\log n / n b_n^s\right) + o_p(n^{-1/2})$$

by (E.11) in Lemma E.3, we have $\left(\hat{D}_{3a}^{(1)}\right)^2 = \left(O_p\left(\log n / n b_n^s\right) + o_p(n^{-1/2})\right) O_p(b_n^{-s})$ and then

$$\frac{K(0)}{n^{1/2} b_n^s} \sum_{s=1}^q \sum_{t=1}^q \hat{D}_{3a}^{(1)} = \left(O_p\left(\sqrt{\log n / n b_n^s}\right) + o_p(n^{-1/4})\right) O_p\left(1/\sqrt{n b_n^{3s}}\right) = o_p(1).$$

Moreover, because $\mathbf{1}_{in} K_{ij} \rho^{(s)}(z_i, \alpha_0) \psi^{(st)}(x_i, \alpha_0)^{-1} \rho^{(t)}(z_i, \alpha_0)$ and $\mathbf{1}_{in} K_{ik} \rho^{(s)}(z_i, \alpha_0) \psi^{(st)}(x_i, \alpha_0)^{-1} \rho^{(t)}(z_k, \alpha_0)$ are uncorrelated for $i \neq j \neq k$,, and for large enough n, $\sup_{x \in \mathcal{X}_n} \psi^{(st)}(x_i, \alpha_0)^{-1} < \infty$, then $\sum_{s=1}^q \sum_{t=1}^q \hat{D}_{3a}^{(2)} = O_p\left(\sqrt{1/n b_n^s}\right)$ so

$$\frac{K(0)}{n^{1/2} b_n^s} \sum_{s=1}^q \sum_{t=1}^q \hat{D}_{3a}^{(2)} = O_p\left(\sqrt{1/n b_n^s}\right) \times O_p\left(\sqrt{1/n b_n^s}\right) = O_p(1/n b_n^s) = o_p(1).$$

Therefore, $b_n^{s/2} \hat{D}_{3a} = o_p(1)$. By similar argument for $\hat{D}_{3a}$, (E.11) in Lemma E.3 and the Cauchy-Schwartz inequality, we have $b_n^{s/2} \hat{D}_{3b} = o_p(1)$ and $b_n^{s/2} \hat{D}_{3c} = o_p(1)$. □

Lemma E.6. *Suppose Assumption A.1-A.9 hold. Then* $b_n^{s/2} \hat{D}_4 \xrightarrow{d} N(0, 2qK^{**} vol(\mathcal{X}_n))$ *under* H_0.

Proof.

$$
\begin{aligned}
\hat{D}_4 &= \sum_{i=1}^{n} \sum_{j=1,j\neq i,}^{n} \sum_{l=1,l\neq j\neq i}^{n} \mathbf{1}_{in} e_{ij} \rho(z_j,\hat{\alpha}_n)' \hat{V}^{-1}(x_i,\hat{\alpha}_n) \rho(z_l,\hat{\alpha}_n) e_{il} \\
&= \sum_{i=1}^{n} \sum_{j=1,j\neq i,}^{n} \sum_{l=1,l\neq j\neq i}^{n} \mathbf{1}_{in} e_{ij} \rho(z_j,\alpha_0)' \hat{V}(x_i,\hat{\alpha}_n)^{-1} \rho(z_l,\alpha_0) e_{il} \\
&\quad + \sum_{i=1}^{n} \sum_{j=1,j\neq i,}^{n} \sum_{l=1,l\neq j\neq i}^{n} \mathbf{1}_{in} e_{ij} \rho(z_j,\alpha_0)' \hat{V}(x_i,\hat{\alpha}_n)^{-1} \left\{\rho(z_l,\hat{\alpha}_n) - \rho(z_l,\alpha_0)\right\} e_{il} \\
&\quad + \sum_{i=1}^{n} \sum_{j=1,j\neq i,}^{n} \sum_{l=1,l\neq j\neq i}^{n} \mathbf{1}_{in} e_{ij} \left\{\rho(z_j,\hat{\alpha}_n) - \rho(z_j,\alpha_0)\right\}' \hat{V}(x_i,\hat{\alpha}_n)^{-1} \rho(z_l,\alpha_0) e_{il} \\
&\quad + \sum_{i=1}^{n} \sum_{j=1,j\neq i,}^{n} \sum_{l=1,l\neq j\neq i}^{n} \mathbf{1}_{in} e_{ij} \left\{\rho(z_j,\hat{\alpha}_n) - \rho(z_j,\alpha_0)\right\}' \hat{V}(x_i,\hat{\alpha}_n)^{-1} \left\{\rho(z_l,\hat{\alpha}_n) - \rho(z_l,\alpha_0)\right\} e_{il} \\
&\equiv \hat{D}_{4a} + \hat{D}_{4b} + \hat{D}_{4c} + \hat{D}_{4d}.
\end{aligned}
$$

We write

$$
\begin{aligned}
\hat{D}_{4a} &= \frac{1}{n^2 b_n^{2s}} \sum_{i=1}^{n} \sum_{j=1,j\neq i,}^{n} \sum_{l=1,l\neq j\neq i}^{n} \mathbf{1}_{in} K_{ij} \rho(z_j,\alpha_0)' H(x_i,\alpha_0)^{-1} \rho(z_l,\alpha_0) K_{il} \\
&\quad + \frac{1}{n^2 b_n^{2s}} \sum_{i=1}^{n} \sum_{j=1,j\neq i,}^{n} \sum_{l=1,l\neq j\neq i}^{n} \mathbf{1}_{in} K_{ij} \rho(z_j,\alpha_0)' \left[\hat{H}(x_i,\alpha_0) - H(x_i,\alpha_0)\right] \rho(z_l,\alpha_0) K_{il} \\
&\equiv \hat{D}_{4a}^{(1)} + \hat{D}_{4a}^{(2)}.
\end{aligned}
$$

For the first term $\hat{D}_{4a}^{(1)}$, let $A_{tj} = \sum_{i=1}^{n} \mathbf{1}_{in} K_{ij} H(x_i,\alpha_0)^{-1} K_{it}$ and $W_{tj} = 2\rho(z_t,\alpha_0)' A_{tj} \rho(z_j,\alpha_0)$, then $\hat{D}_{4a}^{(1)} = \sum_{t=1}^{n-1} \sum_{j=t+1}^{n} W_{tj}$. As in Tripathi and Kitamura (2003), let

$$
\begin{aligned}
G_I &= \sum_{t=1}^{n-1} \sum_{j=t+1}^{n} EW_{tj}^4 \\
G_{II} &= \sum_{t=1}^{n-2} \sum_{j=t+1}^{n-1} \sum_{k=j+1}^{n} \left(EW_{tj}^2 W_{tk}^2 + EW_{jt}^2 W_{jk}^2 + EW_{kt}^2 W_{kj}^2\right)
\end{aligned}
$$

and

$$G_{IV}=\sum_{t=1}^{n-3}\sum_{j=t+1}^{n-2}\sum_{k=j+1}^{n-1}\sum_{l=k+1}^{n}E\left(W_{tj}W_{tk}W_{lj}W_{lk}\right)+E\left(W_{tj}W_{tl}W_{kj}W_{kl}\right)+E\left(W_{tk}W_{tl}W_{jk}W_{jl}\right).$$

Then following similar argument in Lemma A.6 of Tripathi and Kitamura (2003), one can show that $G_I=O(n^4)$, $G_{II}=O(n^5)$ and $G_{IV}=O(n^6)$. Then by the the central limit theorem for generalized quadratic forms in de Jong (1987) (Proposition 3.2) , we have $b_n^{s/2}\hat{D}_{4a}^{(1)}\to N(0,2qK^{**}vol(\mathcal{X}_n))$.

Next, we show that $b_n^{s/2}\hat{D}_{4a}^{(2)}=o_p(1)$. Let $\delta_n(x_i)=\hat{H}(x_i,\alpha_0)-H(x_i,\alpha_0)$. Note that by the uniform consistency of kernel estimators, $\sup_{x\in\mathcal{X}_n}|\delta_n(x)|=O_p\left(\sqrt{\log n/nb_n^s}+b_n^2\right)$. After changing the order of summation, we write

$$\begin{aligned} b_n^{s/2}\hat{D}_{4a}^{(2)} &= \frac{1}{n^2b_n^{3s/2}}\sum_{l=1}^{n}\sum_{j=1,j\neq i,}^{n}\sum_{i=1,i\neq j\neq l}^{n}\mathbf{1}_{in}K_{ij}\rho(z_j,\alpha_0)'\delta_n(x_i)\rho(z_l,\alpha_0)K_{il}, \\ &= \frac{1}{n^2b_n^{3s/2}}\sum_{l=1}^{n}\sum_{j=1,j\neq i,}^{n}\rho(z_j,\alpha_0)'B_{jl}(\delta_n)\rho(z_l,\alpha_0) \\ &= \frac{1}{n^2b_n^{3s/2}}A(\delta_n), \end{aligned} \tag{E.12}$$

where $B_{jl}(\delta_n)\equiv\sum_{i=1,i\neq j\neq l}^{n}\mathbf{1}_{in}K_{ij}\delta_n(x_i)K_{il}$ and $A(\delta_n)\equiv\sum_{l=1}^{n}\sum_{j=1,j\neq i,}^{n}\rho(z_j,\alpha_0)'B_{jl}(\delta_n)\rho(z_l,\alpha_0)$.

Define $B_{jl}^*(\gamma)=\sum_{i=1,i\neq j\neq l}^{n}\mathbf{1}_{in}K_{ij}\gamma K_{il}$, and $A^*(\gamma)=\sum_{l=1}^{n}\sum_{j=1,j\neq l}^{n}\rho(z_j,\alpha_0)B_{jl}^*\rho(z_l,\alpha_0)$. We can see that

$$\{E|A^*(1)|\}^2\leq E\{A^*(1)\}^2=2\sum_{l=1}^{n}\sum_{j=1,j\neq l}^{n}E\{\rho(z_j,\alpha_0)B_{jl}^*(1)\rho(z_l,\alpha_0)\}^2.$$

And we have

$$\begin{aligned} E\{\rho(z_j,\alpha_0)B_{jl}^*(1)\rho(z_l,\alpha_0)\}^2 &= \sum_{i=1,i\neq j\neq l}E\{\mathbf{1}_{in}[\rho(z_j,\alpha_0)]^2[\rho(z_l,\alpha_0)]^2K_{ij}^2K_{il}^2\} \\ &\quad+\sum_{i=1,i\neq j\neq l}\sum_{s=1,s\neq i\neq j\neq l}^{n}E\{\mathbf{1}_{in}[\rho(z_j,\alpha_0)]^2[\rho(z_l,\alpha_0)]^2K_{ij}K_{il}K_{sj}K_{sl}\}. \end{aligned}$$

Note that

$$E\{\mathbf{1}_{in}[\rho(z_j,\alpha_0)]^2[\rho(z_l,\alpha_0)]^2K_{ij}^2K_{il}^2\} = O(b_n^{2s}),$$

and

$$E\{\mathbf{1}_{in}[\rho(z_j,\alpha_0)]^2[\rho(z_l,\alpha_0)]^2K_{ij}K_{il}K_{sj}K_{sl}\} = O(b_n^{3s}).$$

Therefore, we have

$$E\{\rho(z_j,\alpha_0)B_{jl}^*(1)\rho(z_l,\alpha_0)\}^2 = O(n^2b_n^{3s}).$$

Hence $E|A^*(1)| = O(n^2b_n^{3s/2})$. Thus, $A(\delta_n) = o_p(n^2b_n^{3s/2})$ because $\delta_n = o_p(1)$ and $b_n^{s/2}\hat{D}_{4a}^{(2)} = o_p(1)$ by combining with (E.12).

Now let us look at $\hat{D}_{4b}$,

$$\hat{D}_{4b} = \sum_{i=1}^n \sum_{j=1,j\neq i,}^n \sum_{l=1,l\neq j\neq i}^n \mathbf{1}_{in}e_{ij}\rho(z_j,\alpha_0)'\hat{V}(x_i,\hat{\alpha}_n)^{-1}\{\rho(z_l,\hat{\alpha}_n)-\rho(z_l,\alpha_0)\}e_{il}.$$

Let $\Gamma(z_l) = \{\rho(z_l,\hat{\alpha}_n)-\rho(z_l,\alpha_0)\}$ for $l=1,\ldots,n$. Let

$$\begin{aligned}
\hat{D}_{4b} &= \frac{1}{n^2b_n^{2s}}\sum_{i=1}^n \sum_{j=1,j\neq i}^n \sum_{l=1,l\neq j\neq i}^n \mathbf{1}_{in}K_{ij}\rho(z_j,\alpha_0)'H(x_i,\alpha_0)^{-1}\Gamma_n(z_l)K_{il} + o_p(1)\\
&\equiv \frac{1}{n^2b_n^{2s}}\sum_{l=1}^n \sum_{j=1,j\neq l}^n \rho(z_l,\alpha_0)'U_{lj}\Gamma_n(z_j) + o_p(1)\\
&\equiv \frac{1}{n^2b_n^{2s}}\sum_{l=1}^n \sum_{j=1,j\neq l}^n V_{lj}(\Gamma_n) + o_p(1),
\end{aligned}$$

where $U_{lj} = \sum_{i=1,i\neq j\neq l}^n \mathbf{1}_{in}K_{ij}H_n(x_i,\alpha_0)^{-1}K_{il}$ and $V_{lj}(\Gamma_n) = \rho(z_l,\alpha_0)'U_{lj}\Gamma_n(z_j)$. We have $E\{V_{lj}V_{jl}\} = 0$ if $l\neq j$ and $E\{V_{lj}V_{mu}\} = 0$ if $u \neq j$. To show $b_n^{s/2}\hat{D}_{4b} = o_p(1)$, it suffices to show that $\sum_{l=1}^n\sum_{j=1,j\neq l}^n V_{lj}(\Gamma_n) = o_p(n^2b_n^{3s/2})$. Note that since

$$\rho(z_l,\alpha_0)'U_{lj}\Gamma_n(z_j) = \sum_{u=1}^q\sum_{v=1}^q \rho(z_l,\alpha_0)^{(u)}U_{lj}^{(uv)}\Gamma_n^{(v)}(z_j),$$

it suffice to consider $\sum_{l=1}^n\sum_{j=1,j\neq l}^n \rho(z_l,\alpha_0)^{(u)}U_{lj}^{(uv)}\Gamma_n^{(v)}(z_j)$. We already know that $\sup_{(x_i,\alpha)\in\mathcal{X}_n\times\mathcal{A}_n}|\Gamma_n(z_i)| =$

$O_p(\varrho_{s,n})$ by Assumption A.2(on smoothness of ρ) and Assumption 2.1. Consider $V^*_{lj}(\varsigma) = \rho(z_l,\alpha_0)'U_{lj}\varsigma$. Then

$$\begin{aligned} & E\{\rho(z_l,\alpha_0)'U_{lj}\varsigma\}^2 = E\{\rho(z_l,\alpha_0)'A_{lj}\varsigma\varsigma'U_{lj}\rho(z_l,\alpha_0)\} \\ = \;& Etr\left\{A_{lj}\varsigma\varsigma'U_{lj}\rho(z_l,\alpha_0)\rho(z_l,\alpha_0)'\right\} = trE\left\{A_{lj}\varsigma\varsigma'U_{lj}V(x_l,\alpha_0)\right\}. \end{aligned}$$

Because

$$\begin{aligned} & A_{lj}\varsigma\varsigma'U_{lj}V(x_l,\alpha_0) \\ = \;& \left\{\sum_{i=1,i\neq j\neq l}\mathbf{1}_{in}K_{ij}H(x_i,\alpha_0)^{-1}\varsigma\varsigma'K_{il}\right\}\times\left\{\sum_{s=1,s\neq j\neq l}^{n}\mathbf{1}_{sn}K_{sj}H(x_i,\alpha_0)^{-1}V(x_l,\alpha_0)K_{il}\right\}, \end{aligned}$$

we have

$$E\{\rho(z_l,\alpha_0)'U_{lj}\varsigma\}^2 \equiv \sum_{i=1,i\neq j\neq l}^{n} tr(F_1) + \sum_{i=1,i\neq j\neq l}^{n}\sum_{s=1,s\neq i\neq j\neq l} tr(F_2),$$

where

$$I = E\left\{\frac{\mathbf{1}_{in}K_{ij}^2K_{il}^2V(x_i,\alpha_0)^{-1}\varsigma\varsigma'V(x_i,\alpha_0)V(x_l,\alpha_0)}{f^4(x_i)}\right\},$$

$$II = E\left\{\frac{\mathbf{1}_{in}\mathbf{1}_{sn}K_{ij}K_{il}K_{sj}K_{sl}V(x_i,\alpha_0)^{-1}\varsigma\varsigma'V(x_s,\alpha_0)V(x_l,\alpha_0)}{f^2(x_i)f^2(x_s)}\right\}.$$

Note that

$$E[K_{il}^2V(x_l,\alpha_0)|x_i] = b_n^sR(K)V(x_i,\alpha_0)f(x_i) + o_p\left(b_n^{s+2}\right).$$

By iterated expectations and the independence of observations

$$\begin{aligned} F_1 \;&=\; E\left\{\frac{\mathbf{1}_{in}K_{ij}^2V(x_i,\alpha_0)^{-1}\varsigma\varsigma'V(x_i,\alpha_0)}{h^4(x_i)}E[K_{il}^2V(x_l,\alpha_0)|x_i]\right\} \\ &=\; b_n^sR(K)E\left\{\frac{\mathbf{1}_{in}K_{ij}^2V(x_i,\alpha_0)^{-1}\varsigma\varsigma'}{h^3(x_i)}\right\} + b_n^{s+2}E\left\{\frac{\mathbf{1}_{in}K_{ij}^2V(x_i,\alpha_0)^{-1}\varsigma\varsigma'V(x_i,\alpha_0)^{-1}R^{(3)}(x_i)}{h^4(x_i)}\right\} \end{aligned}$$

As $E[K_{ij}^2|x_i] = b_n^sR(K)f(x_i) + o_p\left(b_n^{s+2}\right)$,

$$E\left\{\frac{\mathbf{1}_{in}K_{ij}^2V(x_i,\alpha_0)^{-1}\varsigma\varsigma'}{h^3(x_i)}\right\} = b_n^sR(K)E\left(\frac{\mathbf{1}_{in}V(x_i,\alpha_0)^{-1}\varsigma\varsigma'}{h^2(x_i)}\right)+b_n^{s+2}E\left\{\frac{\mathbf{1}_{in}V(x_i,\alpha_0)^{-1}R^{(4)}(x_i)}{h^3(x_i)}\right\},$$

we have $E\{\rho(z_l,\alpha_0)'U_{lj}\varsigma\}^2 = O(n^2 b_n^{3s})$. The iterated expectations imply that

$$E\{V_{lj}^*(\varsigma)V_{mj}^*(\varsigma)\} = O(n^2 b_n^{4s}).$$

Therefore,

$$\begin{aligned}\Pr\left\{\left|\sum_{l=1}^n \sum_{j=1,j\neq l}^n V_{lj}(\tau_n^{-1/2}\Gamma_n)\right| > M_\varepsilon\right\} &\leq \Pr\left\{\left|\sum_{l=1}^n \sum_{j=1,j\neq l}^n V_{lj}(1)\right| > M_\varepsilon\right\} \\ &\leq E\left|\sum_{l=1}^n \sum_{j=1,j\neq l}^n V_{lj}(1)\right| / M_\varepsilon + o(1),\end{aligned}$$

and

$$\left\{E\left|\sum_{l=1}^n \sum_{j=1,j\neq l}^n V_{lj}(1)\right|\right\}^2 \leq 2\sum_{l=1}^n \sum_{j=1,j\neq l}^n EV_{lj}(1)^2 + \sum_{l=1}^n \sum_{j=1,j\neq l}^n \sum_{m=1,m\neq j\neq l}^n EV_{lj}(1)V_{mj}(1).$$

Because $EV_{lj}(1)^2 = O(n^2 b_n^{3s})$ and $E\{V_{lj}V_{mj}\} = O(n^2 b_n^{4s})$, we get $\left\{E\left|\sum_{l=1}^n \sum_{j=1,j\neq l}^n V_{lj}(1)\right|\right\}^2 = O(n^5 b_n^{4s})$. Then $E\{b_n^{s/2}\hat{D}_{4b}\}^2 = \frac{\varrho_{s,n}^2}{n^4 b_n^{3s}} O(n^5 b_n^{4s}) = O(\varrho_{s,n}^2 n b_n^s) = o(1)$ by conditions in Assumption A.3 (iii). $\square$

Lemma E.7. *Let $\eta_n = n^{-\tau}$ with $\tau \in (0, 1/4]$ and $\mathcal{B}(\eta_n) \equiv \{\alpha \in \mathcal{A}_n : \|\alpha - \alpha_0\| \leq o(\eta_n)\}$. Suppose that Assumptions A.1-A.9 hold. Then*

$$\sup_{\alpha\in\mathcal{B}(\eta_n)} \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in} \|\hat{\lambda}(x_i, \hat{\alpha}_n)\|_E^3 = o_p(\eta_n^3).$$

Proof. First, we show that

$$\sup_{\alpha\in\mathcal{B}(\eta_n)} \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in} \|m(x_i, \alpha)\|_E^3 = o_p(\eta_n^3). \tag{E.13}$$

The argument is similar to that of Corollary A.2 (ii) in Ai and Chen (2003). Because $\mathbf{1}_{in} \leq 1$

and the triangle inequality,

$$
\begin{aligned}
&\sup_{\alpha\in\mathcal{B}(\eta_n)} \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in}\|m(x_i,\alpha)\|_E^3 \\
\leq &\sup_{\alpha\in\mathcal{B}(\eta_n)} \left|\frac{1}{n}\sum_{i=1}^n \left[\|m(x_i,\alpha)\|_E^3 - E\|m(x_i,\alpha)\|_E^3\right]\right| + \sup_{\alpha\in\mathcal{B}(\eta_n)} E\left[\|m(x_i,\alpha)\|_E^3\right].
\end{aligned}
$$

Because of Theorem 2.5.6 in van der Vaart Jon A. Wellner (1996), by Assumption A.2 and A.6, $\{\|m(x_i,\alpha)\|_E^3 : \alpha \in \mathcal{A}_n\}$ is a subset of $\Lambda_c^\gamma(\mathcal{X})$. Thus, it is a Donsker class. Moreover, $m(x_i,\alpha_0) = 0$ and Assumption A.2 implies that as $\|\alpha-\alpha_0\| \to 0$,

$$
E\left[\|m(x_i,\alpha)\|_E^6 - E\|m(x_i,\alpha_0)\|_E^6\right] = E\left[\|m(x_i,\alpha)\|_E^6\right] \leq const. \times \|\alpha-\alpha_0\|^6 \to 0. \qquad \text{(E.14)}
$$

Then $\sup_{\alpha\in\mathcal{B}(\eta_n)} \left|\frac{1}{n}\sum_{i=1}^n [\|m(x_i,\alpha)\|_E^6 - E\|m(x_i,\alpha)\|_E^6]\right| = o_p(n^{-1/2})$ implied by Lemma 1 of Chen *et al.* (2003). Furthermore, by (E.14)

$$
\sup_{\alpha\in\mathcal{B}(\eta_n)} E\left[\|m(x_i,\alpha)\|_E^3\right] \leq \sup_{\alpha\in\mathcal{B}(\eta_n)} \sqrt{E\left[\|m(x_i,\alpha)\|_E^6\right]} \leq o(\eta_n^3),
$$

we have $\sup_{\alpha\in\mathcal{B}(\eta_n)} \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{in}\|m(x_i,\alpha)\|_E^3 = o_p(\eta_n^3)$. Because $\|\hat{\lambda}(x_i,\alpha)\|_E \leq C^{-1}\|\hat{m}(x_i,\alpha)\|_E$ for some constant $C > 0$ by Lemma B.3 (i) in Otsu (2011), the conclusion in (E.13) follows. □

References

Ai, C. and Chen, X. (2003). Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, **71** (6), 1795–1843.

Chen, X., Linton, O. and Van Keilegom, I. (2003). Estimation of semiparametric models when the criterion function is not smooth. *Econometrica*, **71** (5), 1591–1608.

de Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probability Theory and Related Fields*, **75** (2), 261–277.

KITAMURA, Y., TRIPATHI, G. and AHN, H. (2004). Empirical likelihood-based inference in conditional moment restriction models. *Econometrica*, **72** (6), 1667–1714.

OTSU, T. (2011). Empirical likelihood estimation of conditional moment restriction models with unknown functions. *Econometric Theory*, **27** (1), 8–46.

TRIPATHI, G. and KITAMURA, Y. (2003). Testing conditional moment restrictions. *The Annals of Statistics*, **31** (6), 2059–2095.

VAN DER VAART JON A. WELLNER, A. W. (1996). *Weak Convergence and Empirical Processes*. Springer New York.