Technical Proofs for " Smoothing Estimation for Discrete-Valued Time Series" by Z. Cai, Q. Yao and W. Zhang

1 Proofs of Theorem 1

We provide the proof of Theorem 1(ii) only, since Theorem 1(i) may be proved using the standard argument such as the proof of Theorem 3.7 in Lehmann and Casella (1998). The assertion in Remark 2 may be proved in similar and simpler manner.

We use the same notation as in Section 2.3. We present the proof for the case that $\boldsymbol{\beta}_i$ is a scalar, which simplifies the notation without loss of generality. We also assume that the kernel function $K(\cdot)$ has the support $[-1, 1]$.

Proof of Theorem 1(ii). Let $\boldsymbol{\alpha}_i = \dot{\mathbf{b}}(x_i)$, and

$$A = \sum_{t=2}^{n} (1, \delta_n(Y_{t-1} - i)\Big)^T E\Big(\dot{l}(Y_t,\ \mathbf{X}_t,\ \boldsymbol{\beta}_i + \boldsymbol{\alpha}_i \delta_n(Y_{t-1} - i))|\mathbf{X}_t,\ Y_{t-1}\Big) K_{n,h}(Y_{t-1} - i).$$

By the fact that $E\Big(\dot{l}(Y_t,\ \mathbf{X}_t,\ \boldsymbol{\beta}_{Y_{t-1}})|\mathbf{X}_t,\ Y_{t-1}\Big) = 0$, we have

$$\begin{aligned} A \;=\; & \sum_{t=2}^{n} E\left\{\dot{l}(Y_t,\ \mathbf{X}_t,\ \boldsymbol{\beta}_i + \boldsymbol{\alpha}_i \delta_n(Y_{t-1} - i)) - \dot{l}(Y_t,\ \mathbf{X}_t,\ \boldsymbol{\beta}_{Y_{t-1}})|\mathbf{X}_t,\ Y_{t-1}\right\} \times \\ & \Big(1, \delta_n(Y_{t-1} - i)\Big)^T K_{n,h}(Y_{t-1} - i). \end{aligned}$$

Set

$$\xi_t = E\Big\{\dot{l}(Y_t,\ \mathbf{X}_t,\ \boldsymbol{\beta}_i + \boldsymbol{\alpha}_i \delta_n(Y_{t-1} - i)) - \dot{l}(Y_t,\ \mathbf{X}_t,\ \boldsymbol{\beta}_{Y_{t-1}})|\mathbf{X}_t, Y_{t-1}\Big\}, \quad V = \sum_{t=2}^{n} \xi_t K_{n,h}(Y_{t-1} - i).$$

We have

$$\text{var}(V) = \mathbf{I}_1 + \mathbf{I}_2, \tag{A1.1}$$

where

$$\mathbf{I}_1 = (n-1)\text{var}\Big(\xi_2 K_{n,h}(Y_1 - i)\Big), \quad \text{and} \quad \mathbf{I}_2 = 2 \sum_{2 \le k < t \le n} \text{cov}\Big(\xi_k K_{n,h}(Y_{k-1} - i),\ \xi_t K_{n,h}(Y_{t-1} - i)\Big).$$

Condition (A1) together with an application of a Taylor expansion yields

$$
\begin{aligned}
& E\Big(\xi_2^2 K_{n,h}^2(Y_1 - i)\Big) \\
= \;& E\left\{E\Big(\xi_2^2 K_{n,h}^2(Y_1 - i)|Y_1\Big)\right\} \\
= \;& E\left(E\left[E^2\Big\{\ddot{l}(Y_2,\ \mathbf{X}_2,\ \boldsymbol{\beta}_{Y_1})|\mathbf{X}_2, Y_1\Big\}\Big\{\boldsymbol{\beta}_i + \boldsymbol{\alpha}_i\delta_n(Y_1 - i) - \boldsymbol{\beta}_{Y_1}\Big\}^2 K_{n,h}^2(Y_1 - i)|Y_1\right]\right)\Big(1 + o(1)\Big) \\
= \;& E\left(E\left[E^2\Big\{\ddot{l}(Y_2,\ \mathbf{X}_2,\ \boldsymbol{\beta}_{Y_1})|\mathbf{X}_2, Y_1\Big\}|Y_1\right]\Big\{\boldsymbol{\beta}_i + \boldsymbol{\alpha}_i\delta_n(Y_1 - i) - \boldsymbol{\beta}_{Y_1}\Big\}^2 K_{n,h}^2(Y_1 - i)\right)(1 + o(1)) \\
= \;& \frac{1}{4}\ddot{\mathbf{b}}^2(x_i)h^3 g(x_i)\nu_4 E\left[E^2\Big\{\ddot{l}(Y_2,\ \mathbf{X}_2,\ \boldsymbol{\beta}_{Y_1})|\mathbf{X}_2, Y_1\Big\}|Y_1 = i\right](1 + o(1)).
\end{aligned}
$$

Similarly, we can get

$$
E\Big(\xi_2 K_{n,h}(Y_1 - i)\Big) = \frac{1}{2}\ddot{\mathbf{b}}(x_i)h^2 g(x_i)\mu_2\boldsymbol{\Sigma}_i(1 + o(1)). \tag{A1.2}
$$

This leads to

$$
\mathbf{I}_1 = \frac{n-1}{4}\ddot{\mathbf{b}}^2(x_i)h^3 g(x_i)\nu_4 E\left[E^2\Big\{\ddot{l}(Y_2,\ \mathbf{X}_2,\ \boldsymbol{\beta}_{Y_1})|\mathbf{X}_2, Y_1\Big\}|Y_1 = i\right](1 + o(1)). \tag{A1.3}
$$

Now, we consider the term $\mathbf{I}_2$. It is easy to see that

$$
\mathbf{I}_2 = 2\sum_{t=1}^{n-2}(n - 1 - t)\mathrm{cov}\Big(\xi_2 K_{n,h}(Y_1 - i),\ \xi_{t+2}K_{n,h}(Y_{t+1} - i)\Big).
$$

Choose δ, such that

$$
1 - 2/\gamma < \delta < \Big(\beta - \gamma/(\gamma - 2)\Big)(1 - 2/\gamma),
$$

let $d_n = h^{-(1-2/\gamma)/\delta}$,

$$
\mathbf{I}_{2,1} = \sum_{t=1}^{d_n}(n - 1 - t)\mathrm{cov}\Big(\xi_2 K_{n,h}(Y_1 - i),\ \xi_{t+2}K_{n,h}(Y_{t+1} - i)\Big),
$$

and

$$
\mathbf{I}_{2,2} = \sum_{t=d_n+1}^{n-2}(n - 1 - t)\mathrm{cov}\Big(\xi_2 K_{n,h}(Y_1 - i),\ \xi_{t+2}K_{n,h}(Y_{t+1} - i)\Big).
$$

Then $\mathbf{I}_2 = \mathbf{I}_{2,1} + \mathbf{I}_{2,2}$. It holds that for any t,

$$
\mathrm{cov}\Big(\xi_2 K_{n,h}(Y_1 - i),\ \xi_{t+2}K_{n,h}(Y_{t+1} - i)\Big) = O(h^4).
$$

Therefore,

$$(n-1)^{-1}\mathbf{I}_{2,1} = O(h^4 d_n) = O(h^{4-(1-2/\gamma)/\delta}).$$

It follows from the Davydov's inequality (Corollary 1.1 in Bosq, 1996, p.21) that

$$\Big|\mathrm{cov}\Big(\xi_2 K_{n,h}(Y_1-i),\ \xi_{t+2}K_{n,h}(Y_{t+1}-i)\Big)\Big| \le C\Big\{\alpha(t)\Big\}^{1-2/\gamma}\Big\{E|\xi_2 K_{n,h}(Y_1-i)|^\gamma\Big\}^{2/\gamma}.$$

Moreover,

$$\Big\{E|\xi_2 K_{n,h}(Y_1-i)|^\gamma\Big\}^{2/\gamma} \le Ch^4 h^{2/\gamma-2},$$

which leads to

$$(n-1)^{-1}\mathbf{I}_{2,2} \le Ch^4h^{2/\gamma-2}\sum_{t=d_n}^{\infty}\Big\{\alpha(t)\Big\}^{1-2/\gamma} \le Ch^4h^{2/\gamma-2}d_n^{-\delta}\sum_{t=d_n}^{\infty}t^\delta\Big\{\alpha(t)\Big\}^{1-2/\gamma} = o(h^3).$$

Hence, $(n-1)^{-1}\mathbf{I}_2 = o(h^3)$. Together with (A1.1) and (A1.3), this leads to

$$(n-1)^{-1}\mathrm{var}(V) = \frac{1}{4}\ddot{\mathbf{b}}^2(x_i)h^3g(x_i)\nu_4 E\left[E^2\Big\{\ddot{l}(Y_2,\ \mathbf{X}_2,\ \boldsymbol{\beta}_{Y_1})|\mathbf{X}_2, Y_1\Big\}|Y_1=i\right](1+o(1)). \quad \text{(A1.4)}$$

By using the similar arguments, we can show that

$$\begin{aligned}&(n-1)^{-1}\mathrm{var}\left\{\sum_{t=2}^{n}\xi_t\delta_n(Y_{t-1}-i)/hK_{n,h}(Y_{t-1}-i)\right\}\\ =\ &\frac{1}{4}\ddot{\mathbf{b}}^2(x_i)h^3g(x_i)\nu_6 E\left[E^2\Big\{\ddot{l}(Y_2,\ \mathbf{X}_2,\ \boldsymbol{\beta}_{Y_1})|\mathbf{X}_2, Y_1\Big\}|Y_1=i\right](1+o(1)). \quad \text{(A1.5)}\end{aligned}$$

In view of (A1.2), we obtain

$$(n-1)^{-1}\mathbf{G}E(A) = \frac{1}{2}\ddot{\mathbf{b}}(x_i)\boldsymbol{\Sigma}_i(\mu_2,0)^T g(x_i)h^2(1+o(1)),$$

where $\mathbf{G} = \mathrm{diag}(1,\ h^{-1})$. By (A1.4) and (A1.5), we get

$$(n-1)^{-1}\mathbf{G}A = (n-1)^{-1}\mathbf{G}E(A)(1+o_P(1)) = \frac{1}{2}\ddot{\mathbf{b}}(x_i)\boldsymbol{\Sigma}_i(\mu_2,0)^T g(x_i)h^2\Big(1+o_P(1)\Big). \quad \text{(A1.6)}$$

By simple calculation,

$$\ddot{\ell}_n(a,b) = \sum_{t=2}^{n} \ddot{l}\Big(Y_t,\ \mathbf{X}_t,\ a + b\delta_n(Y_{t-1}-i)\Big)\Big(1, \delta_n(Y_{t-1}-i)\Big)^T \Big(1, \delta_n(Y_{t-1}-i)\Big) K_h(\delta_n(Y_{t-1}-i)).$$

Let

$$A_1 = \sum_{t=2}^{n} E\Big\{\ddot{l}(Y_t,\ \mathbf{X}_t,\ \boldsymbol{\beta}_i + \boldsymbol{\alpha}_i\delta_n(Y_{t-1}-i))|\mathbf{X}_t, Y_{t-1}\Big\}(1, \delta_n(Y_{t-1}-i))^T(1, \delta_n(Y_{t-1}-i))K_h(\delta_n(Y_{t-1}-i)).$$

Using the technique in the derivation for (A1.6), we have

$$(n-1)^{-1}\mathbf{G}A_1\mathbf{G} = (n-1)^{-1}\mathbf{G}E(A_1)\mathbf{G}(1+o(1)) = -\boldsymbol{\Sigma}_i \mathrm{diag}(1, \mu_2) g(x_i)(1+o(1)),$$

and

$$(n-1)^{-1}\mathbf{G}\ddot{\ell}_n(\boldsymbol{\beta}_i, \boldsymbol{\alpha}_i)\mathbf{G} - (n-1)^{-1}\mathbf{G}A_1\mathbf{G} = o_P(1),$$

this leads to

$$(n-1)^{-1}\mathbf{G}\ddot{\ell}_n(\boldsymbol{\beta}_i,\ \boldsymbol{\alpha}_i)\mathbf{G} = -\boldsymbol{\Sigma}_i \mathrm{diag}(1, \mu_2) g(x_i)\Big(1 + o_P(1)\Big). \qquad \text{(A1.7)}$$

Therefore, by (A1.6),

$$(1,\ 0)\Big\{\ddot{\ell}_n(\boldsymbol{\beta}_i,\ \boldsymbol{\alpha}_i)\Big\}^{-1} A = -\frac{1}{2}\ddot{\mathbf{b}}(x_i)\mu_2 h^2(1 + o_P(1)). \qquad \text{(A1.8)}$$

It follows by using the same techniques as in the derivation for (A1.4) that

$$\mathrm{var}\Big(\mathbf{G}\Big\{\dot{\ell}_n(\boldsymbol{\beta}_i,\ \boldsymbol{\alpha}_i) - A\Big\}\Big) = (n-1)h^{-1}\boldsymbol{\Sigma}_i \mathrm{diag}(\nu_0, \nu_2) g(x_i)(1+o(1)).$$

By applying to the central limit theorem for kernel regression estimator under α-mixing (see Theorem 3.4 in Bosq, 1996, p.75), we obtain

$$(n-1)^{-1/2}h^{1/2}\Big(\boldsymbol{\Sigma}_i \mathrm{diag}(1, \mu_2) g(x_i)\Big)^{-1}\mathbf{G}\Big\{\dot{\ell}_n(\boldsymbol{\beta}_i,\ \boldsymbol{\alpha}_i) - A\Big\} \to N(0,\ \boldsymbol{\Sigma}_i^{-1}\mathrm{diag}(\nu_0, \nu_2\mu_2^{-2}) g(x_i)^{-1}). \qquad \text{(A1.9)}$$

By following the same line as that used for the proofs of Theorems 3.2-3.10 in Lehmann and

Casella (1998, pp.443-450) and by (A1.7), one has

$$
\begin{aligned}
& \Big\{(n-1)h\Big\}^{1/2}\Big[\hat{\boldsymbol\beta}_i - \boldsymbol\beta_i + (1,\ 0)\Big\{\ddot{\ell}_n(\boldsymbol\beta_i,\ \boldsymbol\alpha_i)\Big\}^{-1} A\Big] \\
= \ & -(1,\ 0)\Big\{(n-1)h\Big\}^{1/2}\Big\{\ddot{\ell}_n(\boldsymbol\beta_i,\ \boldsymbol\alpha_i)\Big\}^{-1}\Big\{\dot{\ell}_n(\boldsymbol\beta_i,\ \boldsymbol\alpha_i) - A\Big\}(1+o_P(1)) \\
= \ & (1,\ 0)(n-1)^{-1/2}h^{1/2}(\boldsymbol\Sigma_i \mathrm{diag}(1,\mu_2)g(x_i))^{-1}\mathbf{G}\Big\{\dot{\ell}_n(\boldsymbol\beta_i,\ \boldsymbol\alpha_i) - A\Big\}(1+o_P(1))
\end{aligned}
$$

This together with (A1.8) and (A1.9) establishes Theorem 1. □

2 Proof of Theorem 2

We provide the proof for Theorem 2 only. All the assertions followed to Theorem 2 may be proved in a similar manner, which is therefore omitted.

We use the same notation as in Section 3.3. Further we always assume that the conditions (B1) — (B4) hold. In order to prove Theorem 2, we need the following lemma.

Lemma 1. *Under the conditions in Theorem 2, we have*

$$
\Big((n-1)\, m^3\, h/\psi_j(x_i)\Big)\, \mathrm{var}(J_1) \ \to\ \nu_0,
$$

where

$$
J_1 = \sum_{t=2}^{n}\Big\{I(Y_t = j) - E(I(Y_t=j)\,|\,Y_{t-1})\Big\}\, w_{as}(t,i)K_{mh}(Y_{t-1}-i), \tag{A2.1}
$$

and

$$
w_{as}(t,i) = (n-1)^{-1}\Big\{1 + \mu_2\,\nu_2^{-1}\,\dot\psi(x_i)\,\psi(x_i)^{-1}\,h\,(Y_{t-1}-i)\,K_{mh}(Y_{t-1}-i)\Big\}^{-1}. \tag{A2.2}
$$

Proof. Let $\varepsilon_{t-1} = I(Y_t = j) - E\{I(Y_t=j)\,|\,Y_{t-1}\}$. Then,

$$
\begin{aligned}
\mathrm{var}(J_1) \ &= \ (n-1)\mathrm{var}\Big(\varepsilon_1\, w_{as}(2,i)\, K_{mh}(Y_1 - i)\Big) \\
&\quad +2\sum_{2\le k<t\le n}\mathrm{cov}\Big(\varepsilon_{k-1}\, w_{as}(k,i)\, K_{mh}(Y_{k-1}-i),\, \varepsilon_{t-1}\, w_{as}(t,i)\, K_{mh}(Y_{t-1}-i)\Big) \\
&\triangleq \ I_1 + I_2,
\end{aligned}
$$

and

$$
\begin{aligned}
I_1 &= \frac{1}{n-1}\sum_{l=1}^{m}\Big(p_{lj}-p_{lj}^2\Big)\Big\{1+\mu_2\,\nu_2^{-1}\,\dot\psi(x_i)\,\psi(x_i)^{-1}h\,(l-i)\,K_{mh}(l-i)\Big\}^{-2}K_{mh}^2(l-i)\,\pi_l \\
&= \Big((n-1)\,m^3\,h\Big)^{-1}\psi_j(x_i)\,\nu_0\,\{1+o(1)\}. \qquad \text{(A2.3)}
\end{aligned}
$$

Choose δ such that

$$1-2/\gamma<\delta<\Big(\beta-\gamma/(\gamma-2)\Big)(1-2/\gamma), \qquad \text{(A2.4)}$$

and let $d_n=m\,h^{-(1-2/\gamma)/\delta}=o\,(m\,h^{-1})$. Set

$$\xi_{t+1}=\varepsilon_{t+1}\Big\{1+\mu_2\,\nu_2^{-1}\,\dot\psi(x_i)\,\psi(x_i)^{-1}\,h\,(Y_{t+1}-i)\,K_{mh}(Y_{t+1}-i)\Big\}^{-1}K_{mh}(Y_{t+1}-i).$$

By stationarity and decomposing I_2 into two terms, we have

$$
\begin{aligned}
I_2 &= 2\,(n-1)^{-2}\sum_{t=1}^{n-2}(n-1-t)\,\mathrm{cov}(\xi_1,\,\xi_{t+1}) \\
&= 2\,(n-1)^{-2}\sum_{t=1}^{d_n}(n-1-t)\,\mathrm{cov}(\xi_1,\,\xi_{t+1})+2\,(n-1)^{-2}\sum_{t=d_n+1}^{n-2}(n-1-t)\,\mathrm{cov}(\xi_1,\,\xi_{t+1}) \\
&\triangleq I_{2,1}+I_{2,2}.
\end{aligned}
$$

For any t, by a simple manipulation, we get

$$\mathrm{cov}(\xi_1,\xi_{t+1})=E(\xi_1\xi_{t+1})=O\left(m^{-4}\right), \qquad \text{(A2.5)}$$

which, coupled with the choice of d_n, implies that

$$I_{2,1}=o\Big(m^{-3}(n-1)^{-1}h^{-1}\Big).$$

Now, we consider contribution of $I_{2,2}$. To this end, by using Davydov's inequality, we obtain

$$|\mathrm{cov}(\xi_1,\,\xi_{t+1})|\le C\{\alpha(t)\}^{1-2/\gamma}\Big\{E|\xi_1|^\gamma\Big\}^{2/\gamma}.$$

Straightforward algebra yields

$$\Big\{E|\xi_1|^\gamma\Big\}^{2/\gamma} \le C m^{-2-2/\gamma} h^{2/\gamma-2},$$

which, in conjunction with the above inequality and Assumption (A1) as well as the choice of d_n, leads to

$$\begin{aligned}(n-1)\, I_{2,2} &\le C m^{-2-2/\gamma}\, h^{2/\gamma-2} \sum_{t=d_n}^{\infty} \Big\{\alpha(t)\Big\}^{1-2/\gamma} \\ &\le C m^{-2-2/\gamma} h^{-1} h^{2/\gamma-1} d_n^{-\delta} \sum_{t=d_n}^{\infty} t^{\delta} \Big\{\alpha(t)\Big\}^{1-2/\gamma} = o(m^{-3} h^{-1}).\end{aligned}$$

This completes the proof of the lemma. □

Proof of Theorem 2.

Let, for $j = 1$ and 2,

$$A_j = \frac{1}{n-1} \sum_{t=2}^{n} (i - Y_{t-1})^j\, K_{mh}^j (Y_{t-1} - i).$$

Similar to calculation of $\mathrm{var}(J_1)$ in Lemma 1, we have

$$\mathrm{var}(A_j) = O(n^{-1} h),$$

which implies that

$$A_1 = E\Big\{(i - Y_1) K_{mh}(Y_1 - i)\Big\} + O_p\left(n^{-1/2} h^{1/2}\right) = -\mu_2\, \dot{\psi}(x_i)\, h^2 + O_p\left(n^{-1/2} h^{1/2} + m^{-1}\right)$$

and

$$A_2 = E\Big\{(i - Y_1)^2 K_{mh}^2 (Y_1 - i)\Big\} + O_P\left(n^{-1/2} h^{1/2}\right) = \nu_2\, \psi(x_i)\, h + O_P\left(n^{-1/2} h^{1/2} + m^{-1}\right).$$

Using the same techniques as in Chen and Hall (1993), we have

$$\begin{aligned}\lambda &= A_2^{-1} A_1 + O_P\left[\Big\{(nh)^{-1/2} + h + (mh)^{-1}\Big\}^2\right] \\ &= -\mu_2 \nu_2^{-1} \dot{\psi}(x_i) \psi(x_i)^{-1} h + O_P\left[\Big\{(nh)^{-1/2} + h + (mh)^{-1}\Big\}^2\right].\end{aligned}$$

Therefore,

$$w_t(i) = w_{as}(t,\, i)\,\{1 + o_P(1)\}.$$

Note that

$$\begin{aligned} & \sum_{t=2}^{n}(I(Y_t = j) - p_{ij})w_t(i)K_{mh}(Y_{t-1} - i) \\ = & \sum_{t=2}^{n}\Big\{I(Y_t = j) - E(I(Y_t = j)|Y_{t-1})\Big\}w_t(i)\,K_{mh}(Y_{t-1} - i) + J_2, \end{aligned}$$

where

$$J_2 = \sum_{t=2}^{n}\Big\{E(I(Y_t = j)|Y_{t-1}) - p_{ij}\Big\}\,w_t(i)\,K_{mh}(Y_{t-1} - i).$$

We have

$$\hat{p}_{ij} - p_{ij} = \frac{\sum_{t=2}^{n}(I(Y_t = j) - p_{ij})\,w_t(i)\,K_{mh}(Y_{t-1} - i)}{\sum_{t=2}^{n} w_t(i)K_{mh}(Y_{t-1} - i)} = (J_1 + J_2)/J_3\,\{1 + o_P(1)\}, \tag{A2.6}$$

where J_1 is defined in (A2.1) and

$$J_3 = \sum_{t=2}^{n} w_{as}(t,\, i)\,K_{mh}(Y_{t-1} - i).$$

By following the same line of the proof in Lemma 1, we obtain

$$J_3 - (n-1)E\Big\{w_{as}(2, i)\,K_{mh}(Y_1 - i)\Big\} = o_P(1),$$

and

$$\begin{aligned} & m\,(n-1)E\Big\{w_{as}(2, i)K_{mh}(Y_1 - i)\Big\} \\ = & m\sum_{j=1}^{m}\Big\{1 + \mu_2\,\nu_2^{-1}\,\dot{\psi}(x_i)\,\psi(x_i)^{-1}\,h\,(j-i)\,K_{mh}(j-i)\Big\}^{-1}\;K_{mh}(j-i)\,\pi_j \\ = & \int_0^1\Big\{1 + \mu_2\,\nu_2^{-1}\,\dot{\psi}(x_i)\,\psi(x_i)^{-1}\,h\,(y - x_i)\,K_h(y - x_i)\Big\}^{-1}\;K_h(y - x_i)\,\psi(y)dy(1 + o(1)) \\ = & \psi(x_i)(1 + o(1)), \end{aligned}$$

so that

$$J_3 = m^{-1}\,\psi(x_i)\,\{1 + o_P(1)\}. \tag{A2.7}$$

By applying the central limit theorem for kernel regression estimator under α-mixing and Lemma 1, we have

$$\sqrt{\frac{(n-1)\,m^3\,h}{\nu_0\,\psi_j(x_i)}}\;J_1 \;\rightarrow\; N(0,\;1). \tag{A2.8}$$

Now, we consider the contribution of J_2. By a Taylor expansion, one has

$$\begin{aligned} m\,(p_{tj}-p_{ij}) &= \frac{\psi_j(x_t)}{\psi(x_t)}-\frac{\psi_j(x_i)}{\psi(x_i)}+O\left(m^{-1}\right) \\ &= \dot{\varphi}_j(x_i)(x_t-x_i)+\frac{1}{2}\,\ddot{\varphi}_j(x_i)(x_t-x_i)^2+O(m^{-1})+o\Big\{(x_t-x_i)^2\Big\}. \end{aligned}$$

By (3.4), we have

$$J_2 \triangleq J_{2,1}\,\{1+o_P(1)\}+O_P(m^{-3})+o_P(m^{-2}h^2),$$

where

$$J_{2,1}=\frac{1}{2m}\,\ddot{\varphi}_j(x_i)\sum_{t=2}^{n}\left(\frac{Y_{t-1}-i}{m}\right)^2\;w_t(i)\,K_{mh}(Y_{t-1}-i).$$

It is easy to see that

$$\begin{aligned} E(J_{2,1}) &= \frac{1}{2\,m}\sum_{t=1}^{m}\Big\{1+\mu_2\nu_2^{-1}\,\dot{\psi}(x_i)\,\psi(x_i)^{-1}\,h\,(t-i)\,K_{mh}(t-i)\Big\}^{-1} \\ &\quad\times\ddot{\varphi}_j(x_i)\,(x_t-x_i)^2\,K_{mh}(t-i)\,\pi_t\,\Big(1+o(1)\Big) \\ &= \frac{h^2}{2m^2}\,\mu_2\,\ddot{\varphi}_j(x_i)\,\psi(x_i)\,\Big(1+o(1)\Big). \end{aligned}$$

Similar to the calculation of $\mathrm{var}(J_1)$, $\mathrm{var}(J_{2,1})=O\Big((n-1)^{-1}m^{-4}h^3\Big)$, so that

$$J_2=\frac{h^2}{2m^2}\,\mu_2\,\ddot{\varphi}_j(x_i)\,\psi(x_i)+o_P\left(m^{-2}h^2\right)+O_P(m^{-3}).$$

This, together with (A2.7) and (A2.8), concludes that

$$\hat{p}_{ij}-p_{ij}=\frac{h^2}{2m}\,\mu_2\,\ddot{\varphi}_j(x_i)+\sqrt{\frac{\nu_0\,\psi_j(x_i)}{(n-1)\,m\,h\,\psi^2(x_i)}}N(0,1)+o_P\Big(((n-1)mh)^{-1/2}+m^{-1}h^2\Big)+O_P(m^{-2}),$$

this leads to the result of Theorem 2. $\square$