Ann. of Appl. Math.
33:1(2017), 63-76

RE-WEIGHTED NADARAYA-WATSON ESTIMATION OF CONDITIONAL DENSITY*†

Yi Wang, Xianzhu Xiong‡

(*College of Math. and Computer Science, Fuzhou University, Fuzhou 350116, Fujian, PR China*)

Abstract

In order to avoid the discussion of equation (1.1), this paper employs the proof method of Liang (2012) to consider the re-weighted Nadaraya-Watson estimation of conditional density. The established results generalize those of De Gooijer and Zerom (2003). In addition, this paper improves the bandwidth condition of Liang (2012).

Keywords re-weighted Nadaraya-Watson estimation; conditional density; bandwidth condition

2000 Mathematics Subject Classification 62N02; 62E20

1 Introduction

Re-weighted Nadaraya-Watson (RNW) method is a weighted version of Nadaraya-Watson (NW). The RNW estimator of the conditional distribution function was proposed by Hall, Wolff and Yao (1999). Later, Cai (2001) applied the RNW method to the estimation of the conditional mean function including the conditional distribution function. The results in Cai (2001) show that the RNW estimation not only possesses the bias of local linear (LL) estimator, but also preserves the property of NW estimator: the estimated values of the conditional mean function are always within the range of the response variable. In the case of estimating a positive quantity such as conditional distribution, the LL method may assign negative weights to certain sample points, and the corresponding LL estimator may produce a negative result in finite samples, thus lead to unreasonable inference. In this case, the RNW estimator works better because it is guaranteed to be nonnegative in finite samples and also has the good bias of the LL estimator. The RNW method has been applied to estimate conditional density (see De Gooijer and Zerom (2003)), to estimate the

*This work was supported by National Natural Science Foundation of China (No.11301084), and Natural Science Foundation of Fujian Province (No.2014J01010).

†Manuscript received July 21, 2016; Revised October 17, 2016

‡Corresponding Author. E-mail: xiongxianzhu2011@gmail.com

volatility function of some diffusion models (see Xu (2010), Hanif, Wang and Lin (2012), Wang, Zhang and Tang (2012), Song, Lin and Wang (2013)), to estimate the conditional variance function (see Xu and Phillips (2011)).

However, in the beginning of proofs of their theorems in Cai (2001) and De Gooijer and Zerom (2003), they used the following equation

$$\sum_{i=1}^{n}\{1+X_{ni}\}Y_{ni}=\{1+o_p(1)\}\sum_{i=1}^{n}Y_{ni}, \quad \text{that is,} \quad \sum_{i=1}^{n}X_{ni}Y_{ni}=o_p(1)\sum_{i=1}^{n}Y_{ni}, \tag{1.1}$$

where $\{(X_{ni}, Y_{ni})|\, 1\le i\le n,\, n\ge 1\}$ is an $\mathbb{R}\times\mathbb{R}$ valued stationary process and $\max\limits_{1\le i\le n}|X_{ni}|=o_p(1)$, (1.1) holds in the case that $\{Y_{ni}\ge 0|\, 1\le i\le n\}$ or $\{Y_{ni}\le 0|\, 1\le i\le n\}$, while (1.1) may be discussed in other cases including $\sum\limits_{i=1}^{n}Y_{ni}=o_p(1)$. Recently, Liang (2012) extended the RNW method of Cai (2001) to the conditional mean function with left-truncated and dependent data, where the author employed another method which can avoid the discussion of equation (1.1) to prove the main result. And in the case of no left truncation, the results are the same to those of Cai (2001). Can the proof method of Liang (2012) be used to consider the RNW estimator of conditional density but avoid the discussion of equation (1.1)?

In this paper, we apply the analysis approach of Liang (2012) to consider the RNW estimator of conditional density. The RNW estimator here is the generalization of that of De Gooijer and Zerom (2003) since it uses different kernel functions and bandwidths in both directions. The established results generalize those of De Gooijer and Zerom (2003). The contributions of this paper are two fold. First, this paper can avoid the discussion of (1.1) in the proof of the results. Second, this paper improves the bandwidth condition in Liang (2012). The rest of the paper is organized as follows. Section 2 introduces the RNW estimator. Assumptions and the main results are stated in Section 3. Section 4 is devoted to proving the main results.

2 Model and RNW Estimator

2.1 Model

Let $\{(X_i, Y_i),\, i\ge 1\}$ be an $\mathbb{R}\times\mathbb{R}$ valued, strictly stationary and α mixing process with a common probability density function $f(\cdot,\cdot)$ as (X, Y). Assume that X admits a marginal density $g(\cdot)$. Of interest is estimating of the conditional density of Y given $X=x$, that is,

$$f(y|x)=\frac{f(x,y)}{g(x)}, \quad y\in\mathbb{R}, \tag{2.1}$$

where $g(x)>0$.

2.2 RNW estimator

From conditions (A1) and (A2) in Section 3, it is easy to verify

$$E\left\{K_{h_n}(X-x)\left[\Lambda_{b_n}(Y-y)-f(y|x)\right]\right\}\to 0, \tag{2.2}$$

where $K_{h_n}(\cdot)=K(\cdot/h_n)/h_n$, $\Lambda_{b_n}(\cdot)=\Lambda(\cdot/b_n)/b_n$, $K(\cdot)$ and $\Lambda(\cdot)$ are two kernel functions defined on $\mathbb{R}$, $0<h_n\to 0$ and $0<b_n\to 0$ as $n\to\infty$. Note that equation (2.2) suggests that $f(y|x)$ can be viewed as a nonparametric regression of $\Lambda_{b_n}(Y_i-y)$ on $\{X_i\}$. In this sense, Rosenblatt (1969) and Fan, Yao and Tong (1996) constructed the NW and LL estimators of $f(y|x)$, respectively. The NW estimator is

$$\widehat{f}_{NW}(y|x)=\sum_{i=1}^{n}\Lambda_{b_n}(Y_i-y)w_i^{NW}(x),\quad w_i^{NW}(x)=\frac{K_{h_n}(X_i-x)}{\sum\limits_{j=1}^{n}K_{h_n}(X_j-x)}. \tag{2.3}$$

The LL estimator is

$$\widehat{f}_{LL}(y|x)=\sum_{i=1}^{n}\Lambda_{b_n}(Y_i-y)w_i^{LL}(x),\quad w_i^{LL}(X)=\frac{K_{h_n}(X_i-x)\left\{T_{n,2}-(X_i-x)T_{n,1}\right\}}{T_{n,0}T_{n,2}-T_{n,1}^2}, \tag{2.4}$$

where $T_{n,k}=\sum\limits_{j=1}^{n}K_{h_n}(X_j-x)(X_j-x)^k$, $k=0,1,2$.

It is easy to see that the LL weights $w_i^{LL}(x)$ satisfy:

$$\sum_{i=1}^{n}w_i^{LL}(x)=1,\quad \sum_{i=1}^{n}(X_i-x)w_i^{LL}(x)=0. \tag{2.5}$$

But for the NW weights $w_i^{NW}(x)$, this moment condition is not fulfilled. Then, similar to the idea of De Gooijer and Zerom (2003), the RNW estimator of $f(y|x)$ is defined as

$$\widehat{f}_{RNW}(y|x)=\sum_{i=1}^{n}\Lambda_{b_n}(Y_i-y)w_i^{RNW}(x),\quad w_i^{RNW}(x)=\frac{p_i(x)K_{h_n}(X_i-x)}{\sum\limits_{j=1}^{n}p_j(x)K_{h_n}(X_j-x)}. \tag{2.6}$$

The sequence of weights $\{p_i(x),\,1\le i\le n\}$ is chosen such that it maximizes $\sum\limits_{i=1}^{n}\log\{p_i(x)\}$ subject to the constraints

$$p_i(x)\ge 0,\quad \sum_{i=1}^{n}p_i(x)=1,\quad \sum_{i=1}^{n}p_i(x)(X_i-x)K_{h_n}(X_i-x)=0. \tag{2.7}$$

By using the Lagrange multiplier used in De Gooijer and Zerom (2003),

$$p_i(x)=\frac{1}{n}\frac{1}{1+\eta(X_i-x)K_{h_n}(X_i-x)},\quad 1\le i\le n, \tag{2.8}$$

where η satisfies:

$$\sum_{i=1}^{n} \frac{(X_i - x)K_{h_n}(X_i - x)}{1 + \eta(X_i - x)K_{h_n}(X_i - x)} = 0. \tag{2.9}$$

3 Assumptions and the Main Result

In the sequel, let C, c_0 and c denote generic finite positive constants, whose values are unimportant and may change from line to line, and let $U(x)$ represent a neighborhood of x. In order to formulate the results, we first provide a list of regularity conditions.

(A1) $K(\cdot)$ and $\Lambda(\cdot)$ are symmetric and bounded density functions with a bounded support $[-1, 1]$.

(A2) (i) $g(x)$ has bounded second order derivative in $U(x)$;

(ii) the second partial derivatives of $f(\cdot|\cdot)$ are continuous in $U(x) \times U(y)$ and $f(y|x) > 0$.

(A3) (i) For all integers $j \geq 1$, the joint density $g_j(\cdot,\cdot)$ of (X_1, X_{j+1}) exists on $\mathbb{R} \times \mathbb{R}$ and satisfies $g_j(s_1, s_2) \leq C$ for $(s_1, s_2) \in U(x) \times U(x)$;

(ii) for all integers $j \geq 1$, the joint density $g_j(\cdot,\cdot,\cdot)$ of (X_1, X_{j+1}, Y_1) exists on $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ and satisfies $g_j(s_1, s_2, t_1) \leq C$ for $(s_1, s_2, t_1) \in U(x) \times U(x) \times U(y)$;

(iii) for all integers $j \geq 1$, the joint density $g_j(\cdot,\cdot,\cdot)$ of (X_1, X_{j+1}, Y_{j+1}) exists on $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ and satisfies $g_j(s_1, s_2, t_2) \leq C$ for $(s_1, s_2, t_2) \in U(x) \times U(x) \times U(y)$;

(iv) for all integers $j \geq 1$, the joint density $g_j(\cdot,\cdot,\cdot,\cdot)$ of $(X_1, X_{j+1}, Y_1, Y_{j+1})$ exists on $\mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ and satisfies $g_j(s_1, s_2, t_1, t_2) \leq C$ for $(s_1, s_2, t_1, t_2) \in U(x) \times U(x) \times U(y) \times U(y)$.

(A4) Assume that $nh_nb_n \to \infty$, and the sequence $\alpha(n)$ satisfies that for positive integers $q = q_n$ there are $q = o((nh_nb_n)^{1/2})$ and $\lim\limits_{n\to\infty}(n(h_nb_n)^{-1})^{1/2}\alpha(q) = 0$.

Remark 3.1 Conditions (A1) and (A2)(i) are used commonly in the literature, see, e.g. Cai (2001), Liang (2012), De Gooijer and Zerom (2003). The property of the bounded support of the kernel functions are often needed in using the dominated convergence theorem in equations (4.23) and (4.36). Condition (A2)(ii) is the same as condition (A2)(ii) in De Gooijer and Zerom (2003).

Remark 3.2 Condition (A3) is mainly technical, which is the same as condition (A3) in Liang and Baek (2016). And it is employed to simplify the calculations of covariances in the proofs, being redundant for independent setting. When $h_n = b_n$, condition (A4) reduces to condition (A4) and (A6) in De Gooijer and Zerom (2003). The role of condition (A4) is to employ Bernstein's big-block and small-block technique to prove asymptotic normality of an α-mixing sequence. In addition, as De Gooijer and Zerom (2003) and Liang and Baek (2016) pointed out, suppose that $h_n = b_n = An^{-\theta}$ $(0 < \theta < 1,\ A > 0)$, $q_n = (nh_n^2/\log n)^{\frac{1}{2}}$ and $\alpha(n) = O(n^{-\lambda})$

($\lambda > 0$), then condition (A6) is satisfied if $\lambda > (1+2\theta)/(1-2\theta)$ (note that λ can be arbitrarily large if $\alpha(n) = O(\rho^n)$ $(0 < \rho < 1)$).

The main result of this paper is presented as follows.

Theorem 3.1 *Let* $\alpha(n) = O(n^{-\lambda})$ *for some* $\lambda > 3$. *Suppose that conditions* (A1)-(A4) *are satisfied. Then*

$$\begin{aligned}\widehat{f}_{RNW}(y|x) - f(y|x) = \frac{1}{2}b_n^2 f^{(0,2)}(y|x)\nabla_{21} + \frac{1}{2}h_n^2 f^{(2,0)}(y|x)\Delta_{21} \\ +o_p(b_n^2 + h_n^2) + O_p((nh_nb_n)^{-\frac{1}{2}}),\end{aligned} \tag{3.1}$$

further

$$\begin{aligned}(nh_nb_n)^{\frac{1}{2}}\Big\{\widehat{f}_{RNW}(y|x) - f(y|x) - \frac{1}{2}h_n^2\Delta_{21}f^{(2,0)}(y|x) \\ -\frac{1}{2}b_n^2\nabla_{21}f^{(0,2)}(y|x) + o_p(h_n^2 + b_n^2)\Big\} \xrightarrow{D} N\big(0, \sigma^2(y|x)\big),\end{aligned} \tag{3.2}$$

where $\sigma^2(y|x) = f(y|x)\Delta_{02}\nabla_{02}/g(x)$, $\Delta_{ij} = \int_{\mathbb{R}} s^i K^j(s)\,\mathrm{d}s$, $\nabla_{ij} = \int_{\mathbb{R}} t^i \Lambda^j(t)\,\mathrm{d}t$, $i, j = 0, 1, 2, \cdots$.

Remark 3.3 If $h_n = b_n$, $K(\cdot) = \Lambda(\cdot)$, and "$nh_n^6 \to c$, $c \neq 0$", then Theorem 3.1 reduces to Theorem 1 in De Gooijer and Zerom (2003).

4 Proof

First we introduce a lemma.

Lemma 4.1 *Let* $\alpha(n) = O(n^{-\lambda})$, $\lambda > 3$. *Set*

$$\begin{aligned}&v_i = (X_i - x)K_{h_n}(X_i - x), \quad 1 \le i \le n,\\ &V_j = \frac{1}{n}\sum_{i=1}^n v_i^j, \quad U_j = \frac{1}{n}\sum_{i=1}^n (X_i - x)^j K_{h_n}(X_i - x), \quad j = 0, 1, 2.\end{aligned}$$

Suppose that conditions (A1), (A2)(i) *and* (A3)(i) *are satisfied. If* $nh_n \to \infty$, *then*

(1) $V_1 = g'(x)\Delta_{21}h_n^2 + O(h_n^3) + O_p((h_n/n)^{\frac{1}{2}})$,
$V_2 = g(x)\Delta_{22}h_n + O(h_n^3) + O_p((h_n/n)^{\frac{1}{2}})$.

(2) $\eta = O_p((nh_n)^{-\frac{1}{2}} + h_n)$, $\quad \max\limits_{1\le i\le n} |\eta v_i| = o_p(1)$.

(3) $U_0 = g(x) + o_p(1)$, $U_2 = g(x)\Delta_{21}h_n^2 + o_p(h_n^2)$.

Remark 4.1 Lemma 4.1(1) and (2) are respectively Lemma 5.1(i) and (ii) of Liang (2012) in the case of no left truncation ($G(y) \equiv 1$, $\theta = 1$, and $\mu(x) = 1$). The proof of Lemma 4.1(3) is similar to that of Lemma 5.1(i) in Liang (2012), and the details are omitted. It should be noted that the proof of Lemma 4.1(2) does not need the condition "$nh_n^{1+r} = O(1)$", which has been used in the proof of Lemma 5.1(ii) in Liang (2012). Therefore, we only prove Lemma 4.1(2).

Proof From equation (2.9),

$$\begin{aligned} 0 &= \Big|\frac{1}{n}\sum_{i=1}^{n}\frac{v_i}{1+\eta v_i}\Big| = \Big|\frac{1}{n}\sum_{i=1}^{n}v_i\Big(1-\frac{\eta v_i}{1+\eta v_i}\Big)\Big| \\ &\geq \frac{|\eta|}{1+\max\limits_{1\leq i\leq n}|\eta v_i|}\frac{1}{n}\sum_{i=1}^{n}v_i^2 - \Big|\frac{1}{n}\sum_{i=1}^{n}v_i\Big| \geq \frac{|\eta|}{1+|\eta|\max\limits_{1\leq i\leq n}|v_i|}V_2 - \big|V_1\big|. \end{aligned} \tag{4.1}$$

Condition (A1) shows that $\max\limits_{1\leq i\leq n}|v_i|\leq C$. Then by equation (4.1),

$$\frac{|\eta|}{1+|\eta|C}V_2 - |V_1| \leq 0.$$

By Lemma 4.1(1),

$$\frac{|\eta|}{1+|\eta|C} = O_p\big((nh_n)^{-\frac{1}{2}} + h_n\big), \tag{4.2}$$

that is, $|\eta|\big(1+O_p((nh_n)^{-\frac{1}{2}}+h_n)\big) = O_p\big((nh_n)^{-\frac{1}{2}}+h_n\big)$, then $\eta = O_p\big((nh_n)^{-\frac{1}{2}}+h_n\big)$. From the fact that $\max\limits_{1\leq i\leq n}|v_i|\leq C$, $\max\limits_{1\leq i\leq n}|\eta v_i| = o_p(1)$.

Proof of Theorem 3.1 Let $m_n(x,y) = E\{\Lambda_{b_n}(Y-y)|\, X=x\}$. From equation (2.6),

$$\widehat{f}_{RNW}(y|x) - f(y|x) = g_n^{-1}(x)(R_n(x)+S_n(x)), \tag{4.3}$$

where

$$\begin{aligned} g_n(x) &= \sum_{i=1}^{n}p_i(x)K_{h_n}(X_i-x), \\ R_n(x) &= \sum_{i=1}^{n}\big(\Lambda_{b_n}(Y_i-y) - m_n(X_i,y)\big)p_i(x)K_{h_n}(X_i-x), \\ S_n(x) &= \sum_{i=1}^{n}\big(m_n(X_i,y) - f(y|x)\big)p_i(x)K_{h_n}(X_i-x). \end{aligned}$$

Step 1 To prove $g_n(x) = g(x)+o_p(1)$, substituting equation (2.8) into $g_n(x)$, we can get

$$\begin{aligned} g_n(x) &= \frac{1}{nh_n}\sum_{i=1}^{n}\frac{1}{1+\eta v_i}K\Big(\frac{X_i-x}{h_n}\Big) \\ &= \frac{1}{nh_n}\sum_{i=1}^{n}K\Big(\frac{X_i-x}{h_n}\Big) - \frac{1}{nh_n}\sum_{i=1}^{n}\frac{\eta v_i}{1+\eta v_i}K\Big(\frac{X_i-x}{h_n}\Big) \\ &:= I_{n1}(x)+I_{n2}(x). \end{aligned} \tag{4.4}$$

By Lemma 4.1(2)(3)

$$|I_{n2}(x)| \le \frac{\max\limits_{1\le i\le n} |\eta v_i|}{1-\max\limits_{1\le i\le n} |\eta v_i|} I_{n1}(x) = o_p(1)O_p(1)U_0 = o_p(1). \tag{4.5}$$

Then,

$$g_n(x) = U_o + o_p(1) = g(x) + o_p(1). \tag{4.6}$$

Step 2 To estimate $S_n(x)$, by conditions (A2)(ii) and (A1), using the second order Taylor expansion about the point (x, y) of the function $f(\cdot|\cdot)$ ($f^{(i,j)}(y|x) := \partial^{i+j} f(y|x)/\partial x^i y^j$), we can get

$$\begin{aligned} m_n(u,y) &= \int \frac{1}{b_n}\Lambda\Big(\frac{s-y}{b_n}\Big) f(s|u)\mathrm{d}s = \int \Lambda(t) f(y+b_n t|u)\mathrm{d}t \\ &= \int \Lambda(t)\Big\{f(y|x) + f^{(1,0)}(y|x)(u-x) + f^{(0,1)}(y|x)b_n t + \frac{1}{2}f^{(2,0)}(y|x)(u-x)^2 \\ &\quad + \frac{1}{2}f^{(0,2)}(y|x)b_n^2 t^2 + f^{(1,1)}(y|x)(u-x)b_n t + o((u-x)^2 + b_n^2 t^2)\Big\}\mathrm{d}t \\ &= f(y|x) + f^{(1,0)}(y|x)(u-x) + \frac{1}{2}f^{(2,0)}(y|x)(u-x)^2 + o((u-x)^2) \\ &\quad + \frac{1}{2}f^{(0,2)}(y|x)\nabla_{21} b_n^2 + o(b_n^2), \quad u \in U(x). \end{aligned} \tag{4.7}$$

Substituting $m_n(X_i, y)$ into $S_n(x)$, one can obtain

$$\begin{aligned} S_n(x) &= \sum_{i=1}^n \big(m_n(X_i,y) - f(y|x)\big)p_i(x)K_{h_n}(X_i - x) \\ &= f^{(1,0)}(y|x)\sum_{i=1}^n (X_i - x)p_i(x)K_{h_n}(X_i - x) \\ &\quad + \frac{1}{2}f^{(2,0)}(y|x)\sum_{i=1}^n \{1+o(1)\}(X_i - x)^2 p_i(x)K_{h_n}(X_i - x) \\ &\quad + \frac{1}{2}f^{(0,2)}(y|x)\nabla_{21} b_n^2 \sum_{i=1}^n \{1+o(1)\}p_i(x)K_{h_n}(X_i - x) \\ &:= S_{n1}(x) + S_{n2}(x) + S_{n3}(x). \end{aligned} \tag{4.8}$$

First we evaluate $S_{n1}(x)$. From equation (2.7), $S_{n1}(x) = 0$. Since $p_i(x)K_{h_n}(X_i - x) \ge 0\,(1 \le i \le n)$,

$$S_{n3}(x) = \frac{1}{2}f^{(0,2)}(y|x)\nabla_{21} b_n^2\{1+o(1)\}\sum_{i=1}^n p_i(x)K_{h_n}(X_i - x),$$

again by the expression of $g_n(x)$ and the fact that $g_n(x) = g(x) + o_p(1)$,

$$S_{n3}(x) = \frac{1}{2}f^{(0,2)}(y|x)g(x)\nabla_{21} b_n^2 + o_p(b_n^2). \tag{4.9}$$

For $S_{n2}(x)$, since $(X_i - x)^2 p_i(x)K_{h_n}(X_i - x) \ge 0\,(1 \le i \le n)$,

$$S_{n2}(x)=\frac{1}{2}f^{(2,0)}(y|x)\{1+o(1)\}\sum_{i=1}^{n}(X_i-x)^2p_i(x)K_{h_n}(X_i-x). \tag{4.10}$$

Substituting equation (2.8) into $\sum\limits_{i=1}^{n}(X_i-x)^2p_i(x)K_{h_n}(X_i-x)$, one can get

$$\begin{aligned}&\sum_{i=1}^{n}(X_i-x)^2p_i(x)K_{h_n}(X_i-x)\\ =&\ \frac{1}{nh_n}\sum_{i=1}^{n}\frac{1}{1+\eta v_i}(X_i-x)^2K\Big(\frac{X_i-x}{h_n}\Big)\\ =&\ \frac{1}{nh_n}\sum_{i=1}^{n}(X_i-x)^2K\Big(\frac{X_i-x}{h_n}\Big)-\frac{1}{nh_n}\sum_{i=1}^{n}\frac{\eta v_i}{1+\eta v_i}(X_i-x)^2K\Big(\frac{X_i-x}{h_n}\Big)\\ :=&\ S_{n21}(x)+S_{n22}(x).\end{aligned} \tag{4.11}$$

By Lemma 4.1(2)(3),

$$S_{n21}(x)=U_2=g(x)\triangle_{21}h_n^2+o_p(h_n^2), \tag{4.12}$$

$$|S_{n22}(x)|\le\frac{\max\limits_{1\le i\le n}|\eta v_i|}{1-\max\limits_{1\le i\le n}|\eta v_i|}U_2=o_p(1)O_p(1)O_p(h_n^2)=o_p(h_n^2). \tag{4.13}$$

It follows from equations (4.10)-(4.13) that

$$S_{n2}(x)=\frac{1}{2}f^{(2,0)}(y|x)g(x)\Delta_{21}h_n^2+o_p(h_n^2). \tag{4.14}$$

Therefore, by the fact that $S_{n1}(x)=0$ and equations (4.8), (4.9) and (4.14), we have

$$S_n(x)=\frac{1}{2}f^{(0,2)}(y|x)g(x)\nabla_{21}b_n^2+\frac{1}{2}f^{(2,0)}(y|x)g(x)\Delta_{21}h_n^2+o_p(b_n^2+h_n^2). \tag{4.15}$$

Step 3 To evaluate $R_n(x)$, noting that

$$p_i(x)=\frac{1}{n}\frac{1}{1+\eta v_i}=\frac{1}{n}\Big\{\sum_{k=0}^{2}(-\eta v_i)^k+\frac{(-\eta v_i)^3}{1+\eta v_i}\Big\},$$

substituting it into $R_n(x)$, we can obtain

$$\begin{aligned}R_n(x)=&\sum_{i=1}^{n}\big(\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)\big)p_i(x)K_{h_n}(X_i-x)\\ =&\frac{1}{nh_n}\sum_{i=1}^{n}\big(\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)\big)K\Big(\frac{X_i-x}{h_n}\Big)\\ &-\frac{\eta}{nh_n}\sum_{i=1}^{n}\big(\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)\big)v_iK\Big(\frac{X_i-x}{h_n}\Big)\end{aligned}$$

$$
\begin{aligned}
&+\frac{\eta^2}{nh_n}\sum_{i=1}^{n}\big(\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)\big)v_i^2K\Big(\frac{X_i-x}{h_n}\Big)\\
&-\frac{\eta^3}{nh_n}\sum_{i=1}^{n}\big(\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)\big)\frac{v_i^3}{1+\eta v_i}K\Big(\frac{X_i-x}{h_n}\Big)\\
:=&\ \frac{1}{nh_n}\left\{R_{n1}(x)-\eta R_{n2}(x)+\eta^2R_{n3}(x)-\eta^3R_{n4}(x)\right\}. \qquad (4.16)
\end{aligned}
$$

From $m_n(X_i,y)=E\{\Lambda_{b_n}(Y_i-y)|X_i\}$ and $v_i=(X_i-x)K_{h_n}(X_i-x)$,

$$E[R_{ni}(x)]=0, \quad i=1,2,3. \qquad (4.17)$$

Let

$$W_i=\Big(\frac{b_n}{h_n}\Big)^{\frac{1}{2}}(\Lambda_{b_n}(Y_i-y)-m_n(X_i,y))K\Big(\frac{X_i-x}{h_n}\Big), \quad 1\le i\le n,$$

then

$$E[W_i]=0, \quad (nh_nb_n)^{\frac{1}{2}}\frac{1}{nh_n}R_{n1}(x)=n^{-\frac{1}{2}}\sum_{i=1}^{n}W_i. \qquad (4.18)$$

It follows from equations (4.31) and (4.32) that

$$Var\Big(n^{-\frac{1}{2}}\sum_{i=1}^{n}W_i\Big)=O(1). \qquad (4.19)$$

It follows from equations (4.17), (4.18) and (4.19) that

$$\frac{1}{nh_n}R_{n1}(x)=O_p((nh_nb_n)^{-\frac{1}{2}}). \qquad (4.20)$$

Similarly,

$$\frac{1}{nh_n}R_{n2}(x)=O_p((nh_nb_n)^{-\frac{1}{2}}), \quad \frac{1}{nh_n}R_{n3}(x)=O_p((nh_nb_n)^{-\frac{1}{2}}). \qquad (4.21)$$

Again by the fact that $\eta=O_p((nh_n)^{-\frac{1}{2}}+h_n)=o_p(1)$,

$$\frac{\eta}{nh_n}R_{n2}(x)=o_p((nh_nb_n)^{-\frac{1}{2}}), \quad \frac{\eta^2}{nh_n}R_{n3}(x)=o_p((nh_nb_n)^{-\frac{1}{2}}). \qquad (4.22)$$

To evaluate $R_{n4}(x)$, equation (4.7) implies that $s\in[-1,1]$, $m_n(x+h_ns,y)\to f(y|x)$, again by conditons (A1) and (A2),

$$
\begin{aligned}
&\frac{1}{nh_n}\sum_{i=1}^{n}E\Big[\,|\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)|\,|v_i^3|K\Big(\frac{X_i-x}{h_n}\Big)\Big]\\
=&\ \frac{1}{nh_n}\sum_{i=1}^{n}E\Big[\,|\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)|\,\frac{|X_i-x|^3}{h_n^3}K^4\Big(\frac{X_i-x}{h_n}\Big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{1}{h_n} E\Big[\Lambda_{b_n}(Y_i-y)\frac{|X_i-x|^3}{h_n^3}K^4\Big(\frac{X_i-x}{h_n}\Big)\Big] \\
&\quad + \frac{1}{h_n} E\Big[\,|m_n(X_i,y)|\,\frac{|X_i-x|^3}{h_n^3}K^4\Big(\frac{X_i-x}{h_n}\Big)\Big] \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \Lambda(t)|s|^3K^4(s)f(x+h_ns,y+b_nt)\mathrm{d}s\mathrm{d}t \\
&\quad + \int_{\mathbb{R}} |m_n(x+h_ns,y)|\,|s|^3K^4(s)g(x+h_ns)\mathrm{d}s \\
&= O(1).
\end{aligned} \tag{4.23}
$$

Then,

$$
\frac{1}{nh_n}\sum_{i=1}^n |\Lambda_{b_n}(Y_i-y)-m_n(X_i,y)|\,|v_i^3|K\Big(\frac{X_i-x}{h_n}\Big)=O_p(1). \tag{4.24}
$$

Again by Lemma 4.1(2),

$$
\begin{aligned}
\frac{|\eta^3|}{nh_n}|R_{n4}(x)| &= \frac{|\eta^3|}{nh_n}\left|\sum_{i=1}^n (\Lambda_{b_n}(Y_i-y)-m_n(X_i,y))\,\frac{v_i^3}{1+\eta v_i}K\Big(\frac{X_i-x}{h_n}\Big)\right| \\
&\le \frac{|\eta|^3}{1-\max\limits_{1\le i\le n}|\eta v_i|}\frac{1}{nh_n}\sum_{i=1}^n |\Lambda_{b_n}(Y_i-y)-m(X_i,y)||v_i^3|K\Big(\frac{X_i-x}{h_n}\Big) \\
&= O_p((nh_n)^{-\frac{3}{2}}+h_n^3) = O_p((nh_n)^{-\frac{3}{2}})+O_p(h_n^3) \\
&= o_p((nh_nb_n)^{-\frac{1}{2}})+o_p(h_n^2).
\end{aligned} \tag{4.25}
$$

It follows from equations (4.3), (4.6), (4.15), (4.16), (4.22), (4.25) and (4.20) that

$$
\begin{aligned}
&\widehat{f}_{RNW}(y|x)-f(y|x) \\
&= \{g^{-1}(x)+o_p(1)\}\Big\{\frac{1}{nh_n}R_{n1}(x)+o_p((nh_nb_n)^{-\frac{1}{2}})+o_p(h_n^2) \\
&\quad +\frac{1}{2}f^{(0,2)}(y|x)g(x)\nabla_{21}b_n^2+\frac{1}{2}f^{(2,0)}(y|x)g(x)\Delta_{21}h_n^2+o_p(b_n^2+h_n^2)\Big\} \\
&= \{g^{-1}(x)+o_p(1)\}\frac{1}{nh_n}R_{n1}(x)+o_p((nh_nb_n)^{-\frac{1}{2}}) \\
&\quad +\frac{1}{2}f^{(0,2)}(y|x)\nabla_{21}b_n^2+\frac{1}{2}f^{(2,0)}(y|x)\Delta_{21}h_n^2+o_p(b_n^2+h_n^2) \\
&= \frac{1}{2}b_n^2f^{(0,2)}(y|x)\nabla_{21}+\frac{1}{2}h_n^2f^{(2,0)}(y|x)\Delta_{21}+O_p((nh_nb_n)^{-\frac{1}{2}})+o_p(b_n^2+h_n^2),
\end{aligned} \tag{4.26}
$$

that is, equation (3.1) follows.

Step 4 To prove equation (3.2), from equation (4.26), it is sufficient to prove

$$(nh_nb_n)^{\frac{1}{2}}\left\{\{g^{-1}(x)+o_p(1)\}\frac{1}{nh_n}R_{n1}(x)+o_p((nh_nb_n)^{-\frac{1}{2}})\right\}\xrightarrow{D}N(0,\sigma^2(y|x)), \quad (4.27)$$

again by equations (4.18), (4.20) and Slutsky's theorem, to prove equation (4.27), it is enough to prove

$$n^{-\frac{1}{2}}\sum_{i=1}^{n}W_i\xrightarrow{D}N(0,g^2(x)\sigma^2(y|x)). \quad (4.28)$$

The proof of equation (4.28) is similar to that of Step 4 in Theorem 3.1 of Liang and Baek (2016). Condition (A4) implies that there exists a sequence $\delta_n\to\infty$ such that $\delta_nq_n=o((nh_nb_n)^{\frac{1}{2}})$, $\delta_n(n(h_nb_n)^{-1})^{\frac{1}{2}}\alpha(q_n)\to 0$. Let $r_n=[\frac{(nh_nb_n)^{\frac{1}{2}}}{\delta_n}]$ and $w_n=[\frac{n}{r_n+q_n}]$. Then

$$\frac{r_n}{(nh_nb_n)^{\frac{1}{2}}}\to 0,\quad \frac{r_n}{n}\to 0,\quad \frac{q_n}{r_n}\to 0,\quad w_n\alpha(q_n)\to 0,\quad \frac{w_nq_n}{n}\to 0. \quad (4.29)$$

Now we employ Bernstein's big-block and small-block procedure. Partition the set $\{1,2,\cdots,n\}$ into $2w_n+1$ subsets with large blocks of size r_n and small blocks of size q_n. Put

$$\xi_{mn}=\sum_{i=k_m}^{k_m+r_n-1}W_i,\quad \xi'_{mn}=\sum_{i=l_m}^{l_m+q_n-1}W_i,\quad \xi''_{mn}=\sum_{i=w_n(r_n+q_n)+1}^{n}W_i, \quad (4.30)$$

where $k_m=(m-1)(r_n+q_n)+1,\ l_m=(m-1)(r_n+q_n)+r_n+1,\ m=1,\cdots,w_n$. Then

$$n^{-\frac{1}{2}}\sum_{i=1}^{n}W_i=n^{-\frac{1}{2}}\left\{\sum_{m=1}^{w_n}\xi_{mn}+\sum_{m=1}^{w_n}\xi'_{mn}+\xi''_{mn}\right\}:=n^{-\frac{1}{2}}\left\{S'_n+S''_n+S'''_n\right\}. \quad (4.31)$$

Note that $E(S'_n)=E(S''_n)=E(S'''_n)=0$, then equation (4.28) follows if one can prove that

$$n^{-1}E(S''_n)^2\to 0,\quad n^{-1}E(S'''_n)^2\to 0,\quad E(n^{-\frac{1}{2}}S'_n)^2\to g^2(x)\sigma^2(y|x); \quad (4.32)$$

$$\left|E\exp\left(it\sum_{m=1}^{w_n}n^{-\frac{1}{2}}\xi_{mn}\right)-\prod_{m=1}^{w_n}E\exp(itn^{-\frac{1}{2}}\xi_{mn})\right|\to 0; \quad (4.33)$$

$$A_n(\varepsilon)=\frac{1}{n}\sum_{m=1}^{w_n}E[\xi_{mn}^2I\left(|\xi_{mn}|>\varepsilon\sqrt{n}\right)]\to 0,\quad \text{for any } \varepsilon>0. \quad (4.34)$$

First we prove equation (4.32). For $n^{-1}E(S''_n)^2$,

$$\begin{aligned}\frac{1}{n}E(S_n'')^2 &= \frac{1}{n}\sum_{m=1}^{w_n}\sum_{i=l_m}^{l_m+q_n-1}E[W_i^2]+\frac{2}{n}\sum_{m=1}^{w_n}\sum_{l_m\le i<j\le l_m+q_n-1}\mathrm{Cov}(W_i,W_j)\\ &\quad+\frac{2}{n}\sum_{1\le i\le j\le w_n}\mathrm{Cov}\left(\xi_{in}',\xi_{jn}'\right)\\ &:= J_{n1}+J_{n2}+J_{n3}. \end{aligned}\tag{4.35}$$

Equation (4.7) implies that $s\in[-1,1]$, $m_n(x+h_ns,y)\to f(y|x)$, again by conditions (A1) and (A2),

$$\begin{aligned}E[W_i^2] &= E\Big[\Big(\frac{b_n}{h_n}\Big)^{\frac{1}{2}}(\Lambda_{b_n}(Y_i-y)-m_n(X_i,y))K\Big(\frac{X_i-x}{h_n}\Big)\Big]^2\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}}K^2(s)\Lambda^2(t)f(x+h_ns,y+b_nt)\mathrm{d}s\mathrm{d}t\\ &\quad+b_n\int_{\mathbb{R}}K^2(s)m_n^2(x+h_ns,y)\,g(x+h_ns)\mathrm{d}s\\ &\quad-2b_n\int_{\mathbb{R}}\int_{\mathbb{R}}K^2(s)\Lambda(t)m_n(x+h_ns,y)f(x+h_ns,y+b_nt)\mathrm{d}s\mathrm{d}t\\ &\to f(x,y)\int_{\mathbb{R}}\int_{\mathbb{R}}K^2(s)\Lambda^2(t)\mathrm{d}s\mathrm{d}t=f(x,y)\Delta_{02}\nabla_{02}. \end{aligned}\tag{4.36}$$

Then, by equations (4.36) and (4.29),

$$J_{n1}=O\Big(\frac{w_nq_n}{n}\Big)=o(1).\tag{4.37}$$

From equation (4.35),

$$|J_{n2}|\le\frac{2}{n}\sum_{1\le i<j\le n}|\mathrm{Cov}(W_i,W_j)|,\quad |J_{n3}|\le\frac{2}{n}\sum_{1\le i<j\le n}|\mathrm{Cov}(W_i,W_j)|.\tag{4.38}$$

Therefore, $J_{n2}=o(1)$ and $J_{n3}=o(1)$ if one can show that

$$\frac{1}{n}\sum_{1\le i<j\le n}|\mathrm{Cov}(W_i,W_j)|\to 0.\tag{4.39}$$

From the fact that $s\in[-1,1]$, $m_n(x+h_ns,y)\to f(y|x)$ and conditions (A1)-(A3), we can obtain by computation, for any $i<j$, $\mathrm{Cov}(W_i,W_j)=O(h_nb_n)$ and $E|W_i(x,y)|^{2\lambda}=O(b_n^{-\lambda+1}h_n^{-\lambda+1})$. On the other hand, it follows from Corollary A.2 in Hall and Heyde (1980, p. 278) that

$$\begin{aligned}|\mathrm{Cov}(W_i,W_j)| &\le C\,[\alpha(j-i)]^{1-\frac{1}{\lambda}}\,(E|W_i|^{2\lambda})^{\frac{1}{\lambda}}\\ &= O(1)[\alpha(j-i)]^{1-\frac{1}{\lambda}}(h_nb_n)^{-(1-\frac{1}{\lambda})}. \end{aligned}\tag{4.40}$$

For $1-1/\lambda<\eta<\lambda-2$, take $c_n=[(h_nb_n)^{-(1-\frac{1}{\lambda})/\eta}]$. Then

$$\begin{aligned}
&\frac{1}{n}\sum_{1\le i<j\le n}|\mathrm{Cov}(W_i,W_j)|\le\frac{1}{n}\sum_{1\le m\le n-1}(n-m)|\mathrm{Cov}(W_i,W_{i+m})|\\
&\le O(1)\Big(\sum_{1\le m\le c_n}+\sum_{c_n+1\le m\le n-1}\Big)\min\big\{\alpha(m)^{1-\frac{1}{\lambda}}(h_nb_n)^{-(1-\frac{1}{\lambda})},h_nb_n\big\}\\
&\le O(1)\Big(c_nh_nb_n+\sum_{c_n+1\le m\le n-1}(h_nb_n)^{-(1-\frac{1}{\lambda})}m^{-\lambda(1-\frac{1}{\lambda})}\Big)\\
&\le O(1)\big(c_nh_nb_n+(h_nb_n)^{-(1-\frac{1}{\lambda})}c_n^{-(\lambda-2)}\big)\to 0, \qquad (4.41)
\end{aligned}$$

that is, equation (4.39) follows. By equations (4.39), (4.38), (4.37) and (4.35), $\frac{1}{n}E(S_n'')^2\to 0$.

For $\frac{1}{n}E(S_n''')^2$, it follows from equations (4.39), (4.36) and (4.29) that

$$\begin{aligned}
\frac{1}{n}E\left(S_n'''\right)^2&=\frac{1}{n}\sum_{i=w_n(r_n+q_n)+1}^{n}E\left(W_i^2\right)+\frac{2}{n}\sum_{w_n(r_n+q_n)+1\le i<j\le n}\mathrm{Cov}(W_i,W_j)\\
&\le C\frac{n-w_n(r_n+q_n)}{n}+\frac{2}{n}\sum_{1\le i<j\le n}|\mathrm{Cov}(W_i,W_j)|\to 0. \qquad (4.42)
\end{aligned}$$

For $\frac{1}{n}E(S_n')^2$, equation (4.29) implies that $w_nr_n/n\to 1$, again by equations (4.36) and (4.39), we obtain

$$\begin{aligned}
\frac{1}{n}E(S_n')^2 &= \frac{1}{n}\sum_{m=1}^{w_n}\sum_{i=k_m}^{k_m+r_n-1}E[W_i^2]+\frac{2}{n}\sum_{m=1}^{w_n}\sum_{k_m\le i<j\le k_m+r_n-1}\mathrm{Cov}(W_i,W_j)\\
&\quad+\frac{2}{n}\sum_{1\le i\le j\le w_n}\mathrm{Cov}(\xi_{in},\xi_{jn})\\
&= \frac{w_nr_n}{n}E[W_i^2]+O\left(\frac{2}{n}\sum_{1\le i<j\le n}|\mathrm{Cov}(W_i,W_j)|\right)\\
&\longrightarrow f(x,y)\Delta_{02}\nabla_{02}=g^2(x)\sigma^2(y|x). \qquad (4.43)
\end{aligned}$$

So equation (4.32) holds.

For equation (4.33), by Lemma 5.1 and equation (4.29),

$$\begin{aligned}
\left|E\exp\left(it\sum_{m=1}^{w_n}n^{-\frac{1}{2}}\xi_{mn}\right)-\prod_{m=1}^{w_n}E\exp\left(itn^{-\frac{1}{2}}\xi_{mn}\right)\right|&\le 16(w_n-1)\alpha(q_n+1)\\
&\le 16w_n\alpha(q_n)\to 0. \qquad (4.44)
\end{aligned}$$

Finally we prove equation (4.34). Equation (4.7) and condition (A1) imply that $m_n(X_i,y)K(\frac{X_i-x}{h_n})=O(1)$, again by condition (A1), $W_i=O(\frac{1}{(h_nb_n)^{1/2}})$. Then,

$\max\limits_{1\le m\le w_n}|\xi_{mn}| = O\big(\frac{r_n}{(h_nb_n)^{1/2}}\big)$. From equation (4.29), $\frac{r_n}{(nh_nb_n)^{1/2}} \to 0$, then $\max\limits_{1\le m\le w_n}\frac{|\xi_{mn}|}{n^{1/2}}$ $= O\big(\frac{r_n}{(nh_nb_n)^{1/2}}\big) = o(1)$, which leads that for large n, $I\left(|\xi_{mn}| > \varepsilon\sqrt{n}\right) = 0$. Therefore, $A_n(\varepsilon) \to 0$.

5 Appendix

Lemma 5.1(Volkonskii (1959)) *Let $Z_1, \cdots, Z_m$ be α-mixing random variables measurable with respect to the σ-algebra $\mathcal{F}_{i_1}^{j_1}, \cdots, \mathcal{F}_{i_m}^{j_m}$ respectively, with $1 \le i_1 < j_1 < \cdots < j_m \le n$, $i_{l+1} - j_l \ge w \ge 1$ and $|Z_j| \le 1$ for $l, j = 1, 2, \cdots, m$. Then*

$$\left|E\left(\prod_{j=1}^m Z_j\right) - \prod_{j=1}^m E(Z_j)\right| \le 16(m-1)\alpha(w),$$

where $\mathcal{F}_a^b = \sigma\{V_i, a \le i \le b\}$ and $\alpha(w)$ is the mixing coefficient.

References

[1] Z. Cai, Weighted Nadaraya–Watson regression estimation, *Statistics* & *Probability Letters*, **51**(2001),307-318.

[2] J.G. De Gooijer, D. Zerom, On conditional density estimation, *Statistica Neerlandica*, **57**(2003),159-176.

[3] P. Hall, R.C. Wolff, Q. Yao, Methods for estimating a conditional distribution function, *Journal of the American Statistical Association*, **94**(1999),154-163.

[4] M. Hanif, H. Wang, Z. Lin, Re-weighted Nadaraya–Watson estimation of jump-diffusion models, *Science China Mathematics*, **55**(2012),1005-1016.

[5] H.Y. Liang, Weighted nonparametric regression estimation with truncated and dependent data, *Journal of Nonparametric Statistic*, **24**(2012),1051-1073.

[6] H.Y. Liang, J.I. Baek, Asymptotic normality of conditional density estimation with left-truncated and dependent data, *Statistical Papers*, **57**(2016),1-20.

[7] M. Rosenblatt, Conditional probability density and regression estimators, *Multivariate analysis II*, (1969),25-31.

[8] Y. Song, Z. Lin, H. Wang, Re-weighted Nadaraya–Watson estimation of second-order jump-diffusion model, *Journal of Statistical Planning and Inference*, **143**(2013),730-744.

[9] Y.Y. Wang, L.X. Zhang, M.T. Tang, Re-weighted functional estimation of second-order diffusion processes, *Metrika*, **75**(2012),1129-1151.

[10] K.L. Xu, Reweighted functional estimation of diffusion models, *Econometric Theory*, **26**(2010),541-563.

[11] K.L. Xu, P.C.B. Phillips, Tilted nonparametric estimation of volatility functions with empirical applications, *Journal of Business* & *Economic Statistics*, **29**(2010),518-528.

(*edited by Mengxin He*)