

UNIFORM CONFIDENCE BANDS FOR LOCAL POLYNOMIAL QUANTILE ESTIMATORS *

CAMILLE SABBAH¹

Abstract. This paper deals with uniform consistency and uniform confidence bands for the quantile function and its derivatives. We describe a kernel local polynomial estimator of quantile function and give uniform consistency. Furthermore, we derive its maximal deviation limit distribution using an approximation in the spirit of Bickel and Rosenblatt [P.J. Bickel and M. Rosenblatt, *Ann. Statist.* **1** (1973) 1071–1095].

Mathematics Subject Classification. 62G08, 62G15.

Received March 3, 2011. Revised June 14, 2012.

1. INTRODUCTION

Consider a scalar dependent random variable Y and an explanatory random variable X . It can be easily shown (Rosenblatt [21]) that if U is uniformly distributed over $[0, 1]$ and independent of X , the dependence between Y and X is expressed by the functional relation

$$Y = Q(U|X), \quad (1.1)$$

which is based on the α -th conditional quantile function $Q(\alpha|x)$ of Y given $X = x$. The representation (1.1) can be viewed as an alternative to the regression model

$$Y = \mathbb{E}[Y|X] + \epsilon. \quad (1.2)$$

However (1.1) is more general than (1.2) which is well defined only if Y has a mean. In (1.2), the conditional mean summarizes the dependence between Y and X , which may be inappropriate in many practical situations. In particular, ϵ may still depend upon X through heteroscedasticity. In this case, the conditional mean fails to fully describe the response of Y to X . This contrasts with (1.1) where the quantile function can be used for such a purpose since U and X are independent. In addition, the representation (1.1) suggests to estimate the derivatives of $Q(\alpha|x)$ with respect to x to obtain the response of Y to an infinitesimal variation of X .

Keywords and phrases. Uniform confidence bands, conditional quantile estimation.

* *This paper was started when the author was at Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie, support from which is gratefully acknowledged. The author thanks the associate editor and two anonymous referees whose careful reading, suggestions, and comments helped to improve the paper. We also would like to thank Professor Emmanuel Guerre for helpful comments and discussions. All remaining errors are our responsibility.*

¹ Laboratoire EQUIPPE, Université Lille Nord de France, Domaine Universitaire du Pont de Bois, BP 60149, 59653 Villeneuve-d'Ascq Cedex, France. camille.sabbah@univ-lille3.fr

The study of conditional quantiles have many other advantages over classical regression. To cite a few, conditional quantile estimators are robust to the presence of outliers, see the seminal paper Huber [11] and Maronna, Martin and Yohai [19] for a general overview about robust statistics. The conditional quantile is also invariant in monotonous transformations, a useful property in *e.g.* censored-data or logistic models, see Huber [12] and Powell [20] among others.

The present work is therefore concerned with some properties of a local polynomial estimator of the conditional quantile function. A first contribution is to establish the limit law of the maximal deviations for a local polynomial estimator of the conditional quantile. A direct use of this limit law is to construct uniform confidence bands for the conditional quantile function and for its derivatives. Uniform confidence bands allow to evaluate the variability in the estimation and at the same time provide a prediction tool. Claeskens and Van Keilegom [3], in the likelihood setup, pointed out that uniform confidence bands are useful to answer graphical queries about the curve's shape. For instance, it can be used in order to choose between a linear quantile regression approach as first developed in [15] and a nonlinear or a nonparametric approach. Uniform confidence bands thus have direct applications in the scope of model specification.

A second contribution of the paper deals with data-driven bandwidths. The coverage performance of the uniform confidence bands of Theorem 2.3 depends upon the asymptotic properties of the estimators of the conditional quantile and of the sparsity function. Indeed, it is known that the bandwidth parameter highly influence the variability of these nonparametric estimators and therefore affects the covering properties of the resulting uniform confidence bands. For instance, by adopting an under-smoothed estimation strategy, we reduce the bias of both estimators and as a consequence, the resulting confidence bands will be more centered around the true curve $x \mapsto Q(\alpha|x)$, but at the same time we increase the variability of both estimators that can lead to a poor coverage performance. See among others Eubanck and Speckman [5] for a discussion about bias correction in the construction of confidence bands. A natural approach to overcome this issue is to use data-driven bandwidths. In Corollary 2.4 below, we state that any random bandwidth converging to 0 at an appropriate rate can be used in the construction of uniform confidence bands. Moreover, when estimating both the conditional quantile and the sparsity function using such plug-in data-driven bandwidths, Guerre and Sabbah ([7], Prop. 2) showed that the resulting estimators achieve the global optimal rates of Stone [22]. While the literature abounds with examples of data-driven bandwidths in the mean regression setup, there are less references concerning such choices for the conditional quantile. Nevertheless, many usual technics as rule-of-thumb, plug-in and cross-validation derived from the mean regression setup are available in the literature to construct adaptive bandwidths for the conditional quantile estimator. See for more details Yu and Jones [26] which proposes a rule-of-thumb for selecting bandwidths for a local linear estimator of the conditional quantile and references therein. Li and Racine [18] proposed to use a data-driven bandwidth initially constructed for the conditional cumulative distribution estimator and to plug it in the estimator of the conditional quantile. Another approach is to consider a cross-validation method. While classical cross-validation technics builds on the L^2 loss function which is very sensible to the presence of outliers in the data, recently Leung [17] proposed to consider a general class of robust cross-validation procedures including *e.g.* the L^1 loss function and studied its asymptotic properties.

Construction of uniform confidence bands have been studied in several contexts. Bickel and Rosenblatt [2] first gave uniform confidence bands for the density of a random variable. In the classical regression setup, see among others Knafl *et al.* [13], Härdle [8] for local constant M -estimation, Sun and Loader [23], Claeskens and Van Keilegom [3] in the likelihood setup and more recently Wang and Yang [25] in the polynomial spline regression estimation context and Härdle and Song [10], among others. Among these references Härdle and Song [10] proposed a closely related to our results by constructing uniform confidence bands for the conditional quantile functions using local constant estimation. We go one step further extending their result by providing confidence bounds not only for the conditional quantile but also for its derivatives. It is known that when the conditional quantile function have a high order of differentiability, its estimator achieves a better bias-variance trade-off. Thus the uniform confidence band of Härdle and Song [10] are suboptimal and can be improved when the conditional quantile function is more than twice differentiable as they assumed.

Now let us describe our setup. From now on we consider independent and identically distributed (i.i.d.) variables (X_i, Y_i) , $i = 1, \dots, n$, having the same distribution than (X, Y) where X and Y are real random variables. Define for α in $(0, 1)$ the conditional quantile function of Y given $X = x$, $Q(\alpha|x) = F^{-1}(\alpha|x)$, where $F(\cdot|x)$ (resp. $f(\cdot|x)$) is the cumulative (resp. density) distribution function of Y given $X = x$. It is well-known that one way of defining the conditional quantile is $Q(\alpha|x) = \arg \min_t \mathbb{E}[\ell_\alpha(Y - t)|X = x]$ where $\ell_\alpha(t) = |t| + (2\alpha - 1)t$. Following this definition, our local polynomial estimator $\widehat{\mathbf{b}}(\alpha, x)$ of order p of $Q(\alpha|x)$ satisfies

$$\widehat{\mathbf{b}}(\alpha, x) = \arg \min_{\mathbf{b} \in \mathbb{R}^{p+1}} \sum_{i=1}^n \ell_\alpha \left(Y_i - \sum_{j=0}^p b_j (X_i - x)^j \right) K \left(\frac{X_i - x}{h} \right),$$

where $\mathbf{b} = (b_0, \dots, b_p)^\top$, h is a bandwidth parameter and $K(\cdot)$ a kernel function. In particular, the first coordinate $\widehat{Q}(\alpha|x) = \widehat{b}_0(\alpha, x)$ of the vector $\widehat{\mathbf{b}}(\alpha, x)$ is an estimator of $Q(\alpha|x)$. Similarly $\widehat{b}_j(\alpha, x)$, the j th coordinate of the vector $\widehat{\mathbf{b}}(\alpha, x)$ is an estimator of $Q_j(\alpha|x) = Q^{(j)}(\alpha|x)/j!$ provided this quantity exists.

Our main result can be expressed as follows. For some deterministic diverging sequences $\{a_n\}$, $\{b_n\}$, and a map $(\alpha, x) \mapsto r_j(\alpha, x)$ defined below in Theorem 2.2 we have for all t and when n diverges

$$\mathbb{P} \left(a_n \left\{ \sup_{x \in \mathcal{X}_0} r_j(\alpha, x)^{-1/2} \left| \widehat{b}_j(\alpha, x) - Q_j(\alpha|x) \right| - b_n \right\} \leq t \right) \rightarrow \exp(-2 \exp(-t)), \tag{1.3}$$

for all $0 \leq j \leq [s]$ and where \mathcal{X}_0 is an inner compact subset of the support of X .

The rest of the paper is organized as follows. Section 2 describes the main toolst for building uniform confidence bands. Furthermore we state the main result and the consequent corollaries. Finally, proofs are gathered in Section 3.

2. MAIN RESULTS

In a first step, we will present a Bahadur representation for the conditional quantile estimator and its derivatives. In a few words, Bahadur [1] first developed a linear representation of the sample quantile with a *small* remainder term. The so-called Bahadur representation is useful to study asymptotic properties of the sample quantile for mostly two reasons. The first one is that it is quite easy to derive asymptotic properties of the conditional quantile estimator since its study is reduced to that of a classical empirical process. The second one is that the remainder term tends to 0 more quickly than the linear term so that this representation well expresses the behavior of the sample quantile.

In order to build uniform confidence bands, we will furthermore use an uniform Bahadur representation. See among others Kong *et al.* [16] which gives an uniform Bahadur representation with respect to the explanatory variable and Guerre and Sabbah [7] which provides a Bahadur representation perhaps better adapted to our context since it is uniform with respect to the level of the quantile, the explanatory variable and the bandwidth parameter. First the uniformity in x allows the construct of uniform confidence bands. Secondly, uniformity in the quantile level allows the construction of an estimator of the sparsity function derived from the conditional quantile estimator. Finally, uniformity in the bandwidth parameter allows the use of data-driven bandwidths.

More specifically, we derive from Guerre and Sabbah ([7], Thm. 2) a new Bahadur representation. For all Assumptions and discussion about, we refer to the Appendix. Let

$$\mathbf{b}(\alpha|x) = (Q(\alpha|x), Q_1(\alpha|x), \dots, Q_{[s]}(\alpha|x), 0, \dots, 0)^\top \in \mathbb{R}^{p+1}, \text{ and where } [s] \text{ is as in Assumption A.4.}$$

Define $\mathbf{N}_p = (\int K(u)u^{i+j}du)_{0 \leq i, j \leq p}$ and $\mathbf{J}(\alpha, x) = 2f(Q(\alpha|x)|x)f(x)\mathbf{N}_p$, where $f(\cdot)$ is the density of X . Furthermore define for any integer number $p \geq [s]$ the vectors in \mathbb{R}^{p+1} , $\mathbf{U}(u) = (1, u, \dots, u^p)^\top$ and $K_{\mathbf{U}}(u) = K(u)\mathbf{U}(u)$. Finally define the $(p + 1) \times (p + 1)$ standardization diagonal matrix $\mathbf{H} = \text{Diag}(h_n^v, v \in \mathbb{N}, v \leq p)$. In the sequel α will be any element of $[\underline{\alpha}, \overline{\alpha}] \subset (0, 1)$.

Lemma 2.1. *Under Assumptions A.1–A.5*

$$(nh_n)^{1/2} \mathbf{H} \left(\widehat{\mathbf{b}}(\alpha, x) - \mathbf{b}(\alpha|x) \right) = \frac{2\mathbf{J}^{-1}(\alpha, x)}{(nh_n)^{1/2}} \sum_{i=1}^n \{ \alpha - \mathbb{I}(Y_i \leq Q(\alpha|X_i)) \} K_{\mathbf{U}} \left(\frac{X_i - x}{h_n} \right) + \mathbf{r}_n(\alpha, x),$$

with $\sup_{x \in \mathcal{X}_0} \|\mathbf{r}_n(\alpha, x)\| = O_{\mathbb{P}} \left((nh_n)^{1/2} h_n^s + h_n \log^{1/2}(n) + (\log(n))^{3/4} / (nh_n)^{1/4} \right)$.

An interesting feature of this result compared with the vast majority of existing results is that the order of the local polynomial estimator may differ from the actual differentiability order $\lfloor s \rfloor$ of $Q(\alpha|x)$.

With the Bahadur’s representation of Theorem 2.1, we are in position to state the limit law of the L^∞ norm of the conditional quantile estimator. Indeed, as explained in the proof section, the study of the limit law of the conditional quantile estimator is reduced to that of the linear term in the RHS in Theorem 2.1 since the remainder term vanishes at an appropriate rate when n diverges.

Our limit distribution result depends upon

$$\begin{aligned} \mathbf{Q}_p &= \left[\int K'(u)^2 u^{i+j} du - \frac{\{i(i-1) + j(j-1)\}}{2} \int K(u) u^{i+j-2} du \right]_{0 \leq i, j \leq p}, \\ \mathbf{T}_p &= \left(\int K^2(u) u^{i+j} du \right)_{0 \leq i, j \leq p}, \quad C_j = \frac{(\mathbf{N}_p^{-1} \mathbf{Q}_p \mathbf{N}_p^{-1})_{j+1, j+1}}{(\mathbf{N}_p^{-1} \mathbf{T}_p \mathbf{N}_p^{-1})_{j+1, j+1}}. \end{aligned}$$

Theorem 2.2. *Assume that Assumptions A.1–A.5 are verified. Then (1.3) holds for all $j = 0, \dots, \lfloor s \rfloor$ with*

$$\begin{aligned} a_n &= (-2 \log(h_n))^{1/2}, \quad b_n = (-2 \log(h_n))^{1/2} + (-2 \log(h_n))^{-1/2} \log(C_j / (2\pi)), \\ r_j(\alpha, x) &= \frac{\alpha(1-\alpha)}{f^2(Q(\alpha|x)|x)f(x)} (\mathbf{N}_p^{-1} \mathbf{T}_p \mathbf{N}_p^{-1})_{j, j}. \end{aligned}$$

Note that $r_j(\alpha, x)$ involves the sparsity function $q(\alpha|x) = f(Q(\alpha|x)|x)^{-1}$ and $f(x)$. Our feasible quantile confidence bound relies on an estimator of $q(\alpha|x)$, $\widehat{q}(\alpha|x)$ which can be found in Guerre and Sabbah ([7], Prop. 3). Let $\widehat{q}(\alpha|x) = \frac{1}{h_n} \int \widehat{Q}(\alpha + h_n t|x) dK_q(t)$, where $K_q(\cdot)$ is a signed measure over \mathbb{R} with compact support \mathcal{K}_q and satisfying

$$\int dK_q(t) = 0, \quad \int t dK_q(t) = 1, \quad \int |dK_q(t)| < \infty.$$

Observe that a Taylor development of $Q(\alpha + h_n t|x)$ in h of order 1 yields that $q(\alpha|x) = \lim_{h_n \rightarrow 0} \frac{1}{h_n} \int Q(\alpha + h_n t|x) dK_q(t)$. This is the idea behind our sparsity function estimation.

Now define $\widehat{f}(x) = \sum_{i=1}^n K((X_i - x)/h_n) / (nh_n)$ and

$$\widehat{\mathbf{V}}(\alpha, x) = \frac{\alpha(1-\alpha)\widehat{q}(\alpha|x)^2}{\widehat{f}(x)} \mathbf{N}_p^{-1} \mathbf{T}_p \mathbf{N}_p^{-1},$$

so that $\left(\widehat{\mathbf{V}}(\alpha, x) \right)_{j, j}$ is an estimator of $r_j(\alpha, x)$. The next Theorem deals with the construction of the uniform confidence bands.

Theorem 2.3. *Under Assumptions A.1–A.5, for any $0 < \lambda < 1$, a $(1 - \lambda)100\%$ confidence band for $x \mapsto Q^{(j)}(\alpha|x)$, $j = 0, \dots, \lfloor s \rfloor$ is given by the collection of all functions belonging to the set of functions \mathcal{Q}_j*

$$\left\{ \mathcal{Q}_j; \sup_{x \in \mathcal{X}_0} \left[\left| j! \widehat{b}_j(\alpha, x) - \mathcal{Q}_j(\alpha, x) \right| \left(\widehat{\mathbf{V}}(\alpha, x) \right)_{j, j}^{-1/2} \right] \leq L_{\lambda, j} \right\},$$

where for all $j = 0, \dots, \lfloor s \rfloor$,

$$L_{\lambda,j} = j! (nh_n^{2j+1})^{-1/2} (-2 \log(h_n)) \left(1 + (-2 \log(h_n))^{-1} \left[a_\lambda + \log \left(C_j^{1/2} / (2\pi) \right) \right] \right),$$

and $a_\lambda = -\log(-\log(1 - \lambda)/2)$.

The next Corollary is concerned with the use of random bandwidths.

Corollary 2.4. *Consider a data-driven bandwidth \hat{h}_n satisfying*

$$\frac{\hat{h}_n}{h_n} = 1 + o_{\mathbb{P}}(1) \quad \text{for some } h_n \text{ as in Assumption A.5,}$$

when n diverges. Then under Assumptions A.1–A.5, Theorem 2.3 remains valid replacing h_n by \hat{h}_n in the estimation of both $\hat{\mathbf{b}}(\alpha, x)$ and $\hat{\mathbf{V}}(\alpha, x)$.

As a concluding remark, note that the convergence in law considered in the present paper is known to be slow (at the rate $O(\log(n)^{-1})$). In order to remedy this inconvenient, several theoretical attempts are available in the literature. Among others Claeskens and Van Keilegom [3] and Härdle *et al.* [9] propose a bootstrap approach in order to ameliorate the rate of convergence of the limit law of their estimators. It is likely that a similar bootstrap procedure can be extended to our setting and lead to a better rate of convergence.

The rest of the paper is devoted to Assumptions and Proofs.

APPENDIX A. ASSUMPTIONS AND PROOFS

Define the lowest integer part $\lfloor s \rfloor$ of s as the unique integer number with $\lfloor s \rfloor < s \leq \lfloor s \rfloor + 1$. For the rest of the paper let $[\underline{\alpha}, \bar{\alpha}]$ be a compact subset of $(0, 1)$. $\|(x, y)\|$ stands for the Euclidean norm and $\mathbb{I}(\cdot)$ for the indicator function. Positive constants are denoted by the generic letter C and may vary from line to line. Our main Assumptions are as follows.

Assumption A.1. The distribution of X has a cumulative distribution function $F(\cdot)$ continuously differentiable over the compact support \mathcal{X} of X and $F'(\cdot) = f(\cdot) \geq C > 0$.

Assumption A.2. The cumulative distribution function $F(\cdot|x)$ of Y given $X = x$ is continuously differentiable over \mathbb{R} with derivative $f(\cdot|x) \geq C > 0$ for all x in \mathcal{X} . The map $(x, y) \in \mathcal{X} \times \mathbb{R} \mapsto f(y|x)$ is Lipschitz continuous.

Assumption A.3. The nonnegative kernel function $K(\cdot)$ has a compact support $\mathcal{K} = [a, b]$ with $\int K(z)dz = 1$. For some $\underline{K} > 0$, $K(\cdot) \geq \underline{K}\mathbb{I}(|\cdot| \leq 1)$. The kernel $K(\cdot)$ is continuously differentiable over \mathcal{K} , and $K(a) = K(b) = 0$.

Assumption A.4. For some real number $s > 1$, $x \mapsto Q(\alpha|x)$ is $\lfloor s + 1 \rfloor$ -times continuously differentiable. The map $x \in \mathcal{X} \mapsto \partial^{\lfloor s+1 \rfloor} Q(\alpha|x) / \partial x^{\lfloor s+1 \rfloor}$ is Hölder continuous with exponent $s - \lfloor s \rfloor$ for all x in \mathcal{X} . The function $\alpha \in [\underline{\alpha}, \bar{\alpha}] \mapsto Q(\alpha|x)$ is continuously differentiable for all α in $[\underline{\alpha}, \bar{\alpha}]$.

Assumption A.5. The order p of the local polynomial estimator satisfies $p \geq \lfloor s \rfloor$ for s as in Assumption A.4. The bandwidth parameter $h = h_n$ is such that $h \rightarrow 0$, $nh \rightarrow \infty$, $\log^5(n)/(nh) \rightarrow 0$, $(nh)^{1/2}h^s \log^{1/2}(n) \rightarrow 0$ when n diverges. Furthermore $\limsup \log(n)/(nh^{s+1})^{1/2} < \infty$ when n diverges.

We now shortly discuss our Assumptions. Assumption A.2 implies in particular that the conditional quantile $Q(\alpha|x)$ is uniquely defined for all α . Assumptions A.1 and A.3 are standard. For $s = 2$ Assumption A.5 is similar to Assumption A2 of Härdle and Song [10] since it allows the bandwidth to belong to the interval $\left[C ((\log^2(n))/n)^{1/3}, \tilde{h}_n \right]$, for any sequence $\tilde{h}_n = o(n \log(n))^{-1/5}$. Finally Assumption A.4 is common in nonparametric statistics.

A.1. Proof of Theorem 2.2

Define $m(z) = \alpha(1 - \alpha)f(z)$, z in \mathcal{X}_0 and observe that there exists $0 < \underline{m} \leq \overline{m} < \infty$ such that

$$\underline{m} \leq \inf_{(\alpha,z) \in [\underline{\alpha}, \overline{\alpha}] \times \mathcal{X}_0} m(z) \leq \sup_{(\alpha,z) \in [\underline{\alpha}, \overline{\alpha}] \times \mathcal{X}_0} m(z) \leq \overline{m},$$

under Assumption A.1. Define for all $0 \leq j \leq p$, $K_j(u) = K(u)u^j$ and

$$Y_{n,j}(x) = \frac{\{\alpha(1 - \alpha)f(x)\}^{-1/2}}{(nh)^{1/2}} \sum_{i=1}^n \{\mathbb{I}(Y_i \leq Q(\alpha|X_i)) - \alpha\} K_j\left(\frac{X_i - x}{h}\right).$$

In a first step we will write the process $Y_{n,j}(x)$ as a particular transformation of the uniform empirical process. In a second step we will compare the process $Y_{n,j}(x)$ with the process

$$Z_j(x) = h^{-1/2} \int K_j\left(\frac{z - x}{h}\right) dW(z),$$

where $W(\cdot)$ is a Wiener process defined on the support of X , and we prove that for all $j = 0, \dots, p$, $Y_{n,j}(x)$ has the same limit distribution than the one of $Z_j(x)$. This will end the proof arguing as in Claeskens and Van Keilegom ([3], proof of Thm. 2.2).

Since $K(\cdot)$ is continuously differentiable and has compact support under Assumption A.3, and since $F(Q(\alpha|X)|X) = \alpha$, we have

$$Y_{n,j}(x) = \frac{\{\alpha(1 - \alpha)f(x)\}^{-1/2}}{(nh)^{1/2}} \sum_{i=1}^n \int \{\alpha - \mathbb{I}(F(Y_i|X_i) \leq \alpha)\} \mathbb{I}(F(X_i) \leq F(z)) dK_j\left(\frac{z - x}{h}\right).$$

Rosenblatt [21] yields that $(F(X_i), F(Y_i|X_i))$ is uniformly distributed over $[0, 1]^2$ so that there exists a sequence of i.i.d. random vectors $(U_i, V_i)_{i \geq 1}$ uniformly distributed on $[0, 1]^2$ such that

$$Y_{n,j}(x) = \frac{\{\alpha(1 - \alpha)f(x)\}^{-1/2}}{(nh)^{1/2}} \sum_{i=1}^n \int \{\alpha - \mathbb{I}(V_i \leq \alpha)\} \mathbb{I}(U_i \leq F(z)) dK_j\left(\frac{z - x}{h}\right).$$

Now let $\mathbb{U}_n(\cdot, \cdot)$ be the bivariate empirical process $\mathbb{U}_n(u, v) = n^{-1/2} \sum_{i=1}^n \{\mathbb{I}(U_i \leq u, V_i \leq v) - uv\}$, (u, v) in $[0, 1]^2$. Then $Y_{n,j}(x)$ can be written w.r.t. $\mathbb{U}_n(\cdot, \cdot)$ as

$$Y_{n,j}(x) = \frac{(\alpha(1 - \alpha)f(x))^{-1/2}}{h^{1/2}} \int \{\alpha \mathbb{U}_n(F(z), 1) - \mathbb{U}_n(F(z), \alpha)\} dK_j\left(\frac{z - x}{h}\right). \tag{A.1}$$

from this last expression, we can combine the Tusnàdy [24] approximation Theorem with the methodology of Bickel and Rosenblatt [2] to achieve the proof.

$Y_{n,j}(x)$ and $Z_j(x)$ have the same limit distribution.

First recall that the result of Tusnàdy ([24], Thm. 1) yields that there exists a sequence of Brownian bridges $(B_n(\cdot, \cdot))_{n \geq 1}$ defined on $[0, 1]^2$ such that

$$\sup_{(u,v) \in [0,1]^2} |\mathbb{U}_n(u, v) - B_n(u, v)| = O_{\mathbb{P}}\left(\frac{\log^2(n)}{n^{1/2}}\right).$$

Now take $Y_{n,j}(x)$ as in (A.1) and observe that

$$\begin{aligned} & \sup_{x \in \mathcal{X}_0} \left| \frac{h^{-1/2}}{m(x)^{1/2}} \int |\alpha (\mathbb{U}_n(F(z), 1) - B_n(F(z), 1)) - \mathbb{U}_n(F(z), \alpha) + B_n(F(z), \alpha)| dK_j \left(\frac{z-x}{h} \right) \right| \\ & \leq \sup_{x \in \mathcal{X}_0} \left| \frac{h^{-1/2}}{m(x)^{1/2}} \int (1 + \alpha) \sup_{(u,v) \in [0,1]^2} |\mathbb{U}_n(u, v) - B_n(u, v)| dK_j \left(\frac{z-x}{h} \right) \right| \\ & \leq \sup_{x \in \mathcal{X}_0} \left| \frac{1}{m(x)^{1/2}} \int dK_j \left(\frac{z-x}{h} \right) \right| \times O_{\mathbb{P}} \left(\frac{\log^2(n)}{(nh)^{1/2}} \right). \end{aligned}$$

The change of variable $z = x + hu$ in the integral above yields that

$$\sup_{x \in \mathcal{X}_0} \left| \frac{1}{m(x)^{1/2}} \int dK_j(u) \right| \leq \left(\inf_{x \in \mathcal{X}_0} m(x)^{1/2} \right)^{-1} \int |K'_j(u)| du.$$

Since $K(\cdot)$ is continuously differentiable over its compact support under Assumption A.3, and since $\inf_{x \in \mathcal{X}_0} m(x) \geq C > 0$, we have that

$$\sup_{x \in \mathcal{X}_0} \left| \frac{1}{m(x)^{1/2}} \int dK_j \left(\frac{z-x}{h} \right) \right| = C.$$

It follows that

$$Y_{n,j}(x) = \frac{h^{-1/2}}{m(x)^{1/2}} \int \{\alpha B_n(F(z), 1) - B_n(F(z), \alpha)\} dK_j \left(\frac{z-x}{h} \right) + \varepsilon_1(\alpha, x), \quad (\text{A.2})$$

with $\sup_{x \in \mathcal{X}_0} |\varepsilon_1(\alpha, x)| = o_{\mathbb{P}}((-\log(h))^{-1/2})$ under Assumption A.5.

Define now a sequence of bivariate Wiener processes $(W_n(\cdot, \cdot))_n \geq 1$ on $[0, 1]^2$ such that $B_n(u, v) = W_n(u, v) - uvW_n(1, 1)$ for all (u, v) in $[0, 1]^2$. Equation (A.2) then yields that

$$\begin{aligned} Y_{n,j}(x) &= \frac{h^{-1/2}}{m(x)^{1/2}} \int (\alpha W_n(F(z), 1) - W_n(F(z), \alpha)) dK_j \left(\frac{z-x}{h} \right) \\ & \quad + W_n(1, 1) \frac{h^{-1/2}}{m(x)^{1/2}} \int (\alpha F(z) - \alpha F(z)) K_j \left(\frac{z-x}{h} \right) + \varepsilon_1(\alpha, x) \\ &= \frac{h^{-1/2}}{m(x)^{1/2}} \int (\alpha W_n(F(z), 1) - W_n(F(z), \alpha)) dK_j \left(\frac{z-x}{h} \right) + \varepsilon_1(\alpha, x) \end{aligned} \quad (\text{A.3})$$

since $W_n(1, 1)$ is almost surely finite. We now rewrite the integral in (A.3). For that observe that since $W_n(u, 0) = W_n(0, t) = 0$ for all (u, t) in $[0, 1]^2$ we have

$$\alpha W_n(F(z), 1) - W_n(F(z), \alpha) = \int \int [\alpha - \mathbb{I}(v \leq \alpha)] \mathbb{I}(u \leq F(z)) dW_n(u, v).$$

Then (A.3) and Fubini's Theorem yield that $Y_{n,j}(x) - \varepsilon_1(\alpha, x)$ can be written as

$$\begin{aligned} & \frac{h^{-1/2}}{m(x)^{1/2}} \int \left(\int \int [\alpha - \mathbb{I}(v \leq \alpha)] \mathbb{I}(u \leq F(z)) dW_n(u, v) \right) dK_j \left(\frac{z-x}{h} \right) \\ &= \frac{h^{-1/2}}{m(x)^{1/2}} \int \int \int [\alpha - \mathbb{I}(v \leq \alpha)] \mathbb{I}(F^{-1}(u) \leq z) dK_j \left(\frac{z-x}{h} \right) dW_n(u, v) \\ &= \frac{h^{-1/2}}{m(x)^{1/2}} \int \int \{\mathbb{I}(v \leq \alpha) - \alpha\} K_j \left(\frac{F^{-1}(u) - x}{h} \right) dW_n(u, v) \end{aligned} \quad (\text{A.4})$$

Observe that $Y_{n,j}(x) - \varepsilon_1(\alpha, x)$ is a centered Gaussian processes. Define $r(\cdot, \cdot)$ the covariance function of the process $Y_{n,j}(x) - \varepsilon_1(\alpha, x)$. Then

$$\begin{aligned} r(x_1, x_2) &= \frac{h^{-1}}{(m(x_1)m(x_2))^{1/2}} \int \int \{\mathbb{I}(v \leq \alpha) - \alpha\}^2 K_j \left(\frac{F(u)^{-1} - x_1}{h} \right) K_j \left(\frac{F(u)^{-1} - x_2}{h} \right) dv du \\ &= \frac{h^{-1}}{(m(x_1)m(x_2))^{1/2}} \int \alpha(1 - \alpha) f(z) K_j \left(\frac{z - x_1}{h} \right) K_j \left(\frac{z - x_2}{h} \right) dz \\ &= \frac{h^{-1}}{(m(x_1)m(x_2))^{1/2}} \int m(z) K_j \left(\frac{z - x_1}{h} \right) K_j \left(\frac{z - x_2}{h} \right) dz = r_1(x_1, x_2), \end{aligned}$$

where $r_1(\cdot, \cdot)$ is the covariance function of the centered Gaussian process

$$x \in \mathcal{X} \mapsto h^{-1/2} \int \left(\frac{m(z)}{m(x)} \right)^{1/2} K_j \left(\frac{z - x}{h} \right) dW(z),$$

and where $W(\cdot)$ is a Wiener process defined on \mathcal{X} . It then follows from (A.4) that $Y_{n,j}(x) - \varepsilon_1(\alpha, x)$ have the same distribution than

$$x \in \mathcal{X} \mapsto h^{-1/2} \int \left(\frac{m(z)}{m(x)} \right)^{1/2} K_j \left(\frac{z - x}{h} \right) dW(z).$$

To achieve the proof, it remains to show that

$$A = \sup_{x \in \mathcal{X}} \left| h^{-1/2} \int \left(\left[\frac{m(z)}{m(x)} \right]^{1/2} - 1 \right) K_j \left(\frac{z - x}{h} \right) dW(z) \right| = o_{\mathbb{P}}((-\log(h))^{1/2}). \quad (\text{A.5})$$

The change of variable $z = x + hu$ and integration by parts in (A.5) under Assumptions A.1 and A.3 yield that

$$\begin{aligned} & \left| h^{-1/2} \int \left(\left[\frac{m(z)}{m(x)} \right]^{1/2} - 1 \right) K_j \left(\frac{z - x}{h} \right) dW(z) \right| \\ & \leq \left| h^{-1/2} \int W(x + hu) \left(\left[\frac{f(x + hu)}{f(x)} \right]^{1/2} - 1 \right) K_j'(u) du + \frac{h^{1/2}}{2} \int W(x + hu) \frac{f'(x + hu) K_j(u)}{(f(x) f(x + hu))^{1/2}} du \right|. \end{aligned}$$

Since $f(\cdot)$ is continuously differentiable, bounded away from 0 under Assumption X and $\sup_{u \in \mathcal{X}_0} |W(u)| = O_{\mathbb{P}}(1)$, Assumption A.3 yields that the second integral in the inequality above is a $O_{\mathbb{P}}(h^{1/2})$ uniformly in x in \mathcal{X}_0 . To control the first integral in the inequality above, observe continuous differentiability of $f(\cdot)$ under Assumption A.1 and compactness of \mathcal{X}_0 yield that $\sup_{x \in \mathcal{X}_0} |[f(x + hu)/f(x)]^{1/2} - 1| \leq Ch|u|$. Hence $A \leq |\int CK_j'(u) |u| du| O_{\mathbb{P}}(h^{1/2})$, since $\sup_{x \in \mathcal{X}_0} |W(x)| = O_{\mathbb{P}}(1)$, which gives the result under Assumption A.3.

A.2. Proof of Theorem 2.3

The Theorem is proved if we show that

$$\sup_{x \in \mathcal{X}_0} \left\| \widehat{\mathbf{V}}(\alpha, x) - \frac{\alpha(1 - \alpha)q(\alpha|x)^2}{f(x)} \mathbf{N}_p^{-1} \mathbf{T}_p \mathbf{N}_p^{-1} \right\| = o_{\mathbb{P}}(\log(n))^{-1/2}.$$

First observe that Einmahl and Mason ([4], Thm. 1) under Assumptions A.1, A.3 and A.5 yields that

$$\sup_{x \in \mathcal{X}_0} \left| f(x) - \frac{1}{nh} \sum_{i=1}^n K \left(\frac{X_i - x}{h} \right) \right| = O_{\mathbb{P}} \left(\frac{\log^{1/2}(n)}{(nh)^{1/2}} + h \right).$$

Then since $\inf_{x \in \mathcal{X}_0} f(x) \geq C > 0$ under Assumption A.1 and $\inf_{x \in \mathcal{X}_0} q(\alpha|x) \geq C > 0$ under Assumption A.2, the Theorem is proved if we show that

$$\sup_{x \in \mathcal{X}_0} |\widehat{q}(\alpha|x) - q(\alpha|x)| = o_{\mathbb{P}}(\log(n))^{-1/2}.$$

Guerre and Sabbah ([7], Lems. A.2, A.3 and Thms. 1 and 2) under Assumptions A.1, A.2, A.3, A.4, and A.5 yields that

$$\sup_{(\alpha,x) \in [\underline{\alpha}, \overline{\alpha}] \times \mathcal{X}_0} \left| \widehat{Q}(\alpha|x) - Q(\alpha|x) \right| = O_{\mathbb{P}} \left(\frac{\log^{1/2}(n)}{(nh)^{1/2}} + h^{s+1} \right)$$

Now recall that $\widehat{q}(\alpha|x) = \frac{1}{h} \int \widehat{Q}(\alpha + ht|x) dK_q(t)$. Then

$$\begin{aligned} \sup_{x \in \mathcal{X}_0} \left| \widehat{q}(\alpha|x) - \frac{1}{h} \int Q(\alpha + ht|x) dK_q(t) \right| &\leq \frac{1}{h} \sup_{(\alpha,x) \in [\underline{\alpha}, \overline{\alpha}] \times \mathcal{X}_0} \left| \widehat{Q}(\alpha|x) - Q(\alpha|x) \right| \int |dK_q(t)| \\ &= O_{\mathbb{P}} \left(\frac{\log^{1/2}(n)}{h(nh)^{1/2}} + h^s \right). \end{aligned}$$

Now since

$$\sup_{(\alpha,x) \in [\underline{\alpha}, \overline{\alpha}] \times \mathcal{X}_0} \sup_{t \in \mathcal{K}_q} |Q(\alpha + ht) - Q(\alpha|x) - htq(\alpha|x)| = O(h),$$

under Assumptions A.4 and A.5, and since $\int dK_q(t) = 0$ and $\int t dK_q(t) = 1$, we have

$$\sup_{x \in \mathcal{X}_0} |\widehat{q}(\alpha|x) - q(\alpha|x)| = O_{\mathbb{P}} \left(\frac{\log^{1/2}(n)}{h(nh)^{1/2}} + h^s + h \right) = o_{\mathbb{P}}(\log(n))^{-1/2}, \text{ under Assumption A.5.}$$

A.2.1. Proof of Corollary 2.4

The proof is a direct consequence of Guerre and Sabbah ([7], Prop. 2), so it is omitted.

A.3. Proof of Lemma 2.1

The vector $\widehat{\mathbf{b}}(\alpha, x)$ is in a first step an estimator of

$$\mathbf{b}^*(\alpha, x) = \arg \min_{\mathbf{b} \in \mathbb{R}^{p+1}} \mathbb{E} \left[\ell_{\alpha} (Y - \mathbf{U}(X - x)^{\top} \mathbf{b}) K \left(\frac{X - x}{h} \right) \right].$$

Under Assumptions A.1–A.5 and provided h is small enough we have by Guerre and Sabbah ([7], Thm. 1),

$$\sup_{x \in \mathcal{X}_0} \left| \frac{\mathbf{H}(\mathbf{b}^*(\alpha, x) - \mathbf{b}(\alpha|x))}{h^s} \right| \leq C. \tag{A.6}$$

We now present the Bahadur representation given in Guerre and Sabbah [7]. Define $Q^*(X_i; \alpha, x) = \mathbf{U}(X_i - x)^{\top} \mathbf{b}^*(\alpha, x)$ and

$$\begin{aligned} \mathbf{S}_n(\alpha, x) &= \frac{2}{(nh)^{1/2}} \sum_{i=1}^n \{ \mathbb{I}(Y_i \leq Q^*(X_i; \alpha, x)) - \alpha \} K_{\mathbf{U}} \left(\frac{X_i - x}{h} \right), \\ \mathbf{J}_n(\alpha, x) &= \frac{2}{nh} \sum_{i=1}^n f(Q^*(X_i; \alpha, x)|X_i) K_{\mathbf{U}} \left(\frac{X_i - x}{h} \right) \mathbf{U} \left(\frac{X_i - x}{h} \right)^{\top}, \\ \mathbf{E}_n(\alpha, x) &= (nh)^{1/2} \mathbf{H} \left(\widehat{\mathbf{b}}(\alpha, x) - \mathbf{b}^*(\alpha, x) \right) + \frac{\mathbf{S}_n(\alpha, x)}{\mathbf{J}_n(\alpha, x)}. \end{aligned}$$

Under Assumptions A.1–A.5, Guerre and Sabbah ([7], Thm. 2) yields that

$$\sup_{x \in \mathcal{X}_0} \|\mathbf{E}_n(\alpha, x)\| = O_{\mathbb{P}} \left(\frac{\log^3(n)}{nh} \right)^{1/4}. \tag{A.7}$$

Lemma A.6. *Under Assumptions A.1–A.3, A.5 and uniformly in x in \mathcal{X}_0 ,*

$$\left\| \mathbf{S}_n(\alpha, x) - \frac{2}{(nh)^{1/2}} \sum_{i=1}^n \mathbb{I}(Y_i \leq Q(\alpha|X_i)) K_{\mathbf{U}} \left(\frac{X_i - x}{h} \right) \right\| = O_{\mathbb{P}} \left((nh)^{1/2} h^s \right),$$

and $\sup_{x \in \mathcal{X}_0} \|\mathbf{J}(\alpha, x) - \mathbf{J}_n(\alpha, x)\| = O_{\mathbb{P}} \left(h + (\log(n))^{1/2}/(nh)^{1/2} \right)$.

We now return to the Proof of Lemma 2.1. We consider the Bahadur representation in equation (A.7) so that $(nh)^{1/2} \mathbf{H}(\widehat{\mathbf{b}}(\alpha, x) - \mathbf{b}(\alpha, x))$ can be written as

$$\frac{2\mathbf{J}(\alpha, x)^{-1}}{(nh)^{1/2}} \sum_{i=1}^n \{ \alpha - \mathbb{I}(Y_i \leq Q(\alpha|X_i)) \} K_{\mathbf{U}} \left(\frac{X_i - x}{h} \right) + \mathbf{r}_n(\alpha, x),$$

with $\mathbf{r}_n(\alpha, x) = \sum_{j=1}^4 \mathbf{r}_{n,j}(\alpha, x)$, and where

$$\begin{aligned} \mathbf{J}(\alpha, x) \mathbf{r}_{n,1}(\alpha, x) &= \frac{1}{(nh)^{1/2}} \sum_{i=1}^n \{ \mathbb{I}(Y_i \leq Q(\alpha|X_i)) - \alpha \} K_{\mathbf{U}} \left(\frac{X_i - x}{h} \right) - \mathbf{S}_n(\alpha, x), \\ \mathbf{r}_{n,2}(\alpha, x) &= (nh)^{1/2} \mathbf{H}(\mathbf{b}^*(\alpha, x) - \mathbf{b}(\alpha|x)), \quad \mathbf{r}_{n,3}(\alpha, x) = \frac{\mathbf{J}_n(\alpha, x) - \mathbf{J}(\alpha, x)}{\mathbf{J}(\alpha, x) \mathbf{J}_n(\alpha, x)} \mathbf{S}_n(\alpha, x), \end{aligned}$$

and $\mathbf{r}_{n,4}(\alpha, x) = \mathbf{E}_n(\alpha, x)$. We now bound the $\mathbf{r}_{n,j}(\alpha, x)$ for $j = 1, 2, 3, 4$, uniformly in x in \mathcal{X}_0 showing that $\max_{1 \leq j \leq 4} \sup_{x \in \mathcal{X}_0} \|\mathbf{r}_{n,j}(\alpha, x)\| = o_{\mathbb{P}}(\log(n))^{-1/2}$. Lemma A.6 and the fact $\inf_{x \in \mathcal{X}_0} \|\mathbf{J}(\alpha, x)\| \geq C > 0$ yield that $\mathbf{r}_{n,1}(\alpha, x)$ is a $O_{\mathbb{P}}((nh)^{1/2} h^s)$ uniformly in x in \mathcal{X}_0 . Equation (A.6) yields that $\mathbf{r}_{n,2}(\alpha, x)$ is a $O(h^s)$ uniformly in x in \mathcal{X}_0 . Now observe that Guerre and Sabbah ([7], Lem. A.3) together with the fact that Guerre and Sabbah ([7], Lem. A.2) yields that $\inf_{x \in \mathcal{X}_0} \|\mathbf{J}_n(\alpha, x)\| \geq C > 0$ with a probability that can be arbitrarily large yield that $\mathbf{r}_{n,3}(\alpha, x)$ is a $O_{\mathbb{P}}(\log(n)/(nh)^{1/2} + h \log^{1/2}(n))$ uniformly in x in \mathcal{X}_0 . Finally (A.7) yields that $\sup_{x \in \mathcal{X}_0} \|\mathbf{E}_n(\alpha, x)\| = O_{\mathbb{P}}(\log^3(n)/(nh))^{1/4}$. Then

$$(nh)^{1/2} \mathbf{H} \left(\widehat{\mathbf{b}}(\alpha, x) - \mathbf{b}(\alpha|x) \right) = -\frac{\mathbf{J}(\alpha, x)^{-1}}{(nh)^{1/2}} \sum_{i=1}^n \{ \mathbb{I}(Y_i \leq Q(\alpha|X_i)) - \alpha \} + \mathbf{r}_n(\alpha, x),$$

with

$$\sup_{x \in \mathcal{X}_0} \|\mathbf{r}_n(\alpha, x)\| = O_{\mathbb{P}} \left((nh)^{1/2} h^s + \frac{\log(n)}{(nh)^{1/2}} + h \log^{1/2}(n) + \frac{\log^{3/4}(n)}{(nh)^{1/4}} \right).$$

The result then follows since $(\log(n))/(nh)^{1/2} = o \left((\log^{3/4}(n))/(nh)^{1/4} \right)$ under Assumption A.5.

A.4. Proof of Lemma A.6

For simplicity of notations, define $Q^*(u) = Q^*(u; \alpha, x)$ and $Q(u) = Q(\alpha|u)$. A Taylor expansion of $Q(\cdot)$ in a neighborhood of x can be written as

$$Q(x + hz) = \mathbf{U}(hz) \mathbf{b}(\alpha, x) + R_Q(\alpha, x, z),$$

where $\mathbf{b}(\alpha, x)$ is as in Lemma 2.1. This gives that

$$\begin{aligned} \sup_{(x,z) \in \mathcal{X}_0 \times \mathcal{K}} |Q^*(x + hz) - Q(x + hz)| &= \sup_{(x,z) \in \mathcal{X}_0 \times \mathcal{K}} |\mathbf{U}(hz)\mathbf{b}^*(\alpha, x) - \mathbf{U}(hz)\mathbf{b}(\alpha, x) - R_Q(\alpha, x)| \\ &\leq \sup_{z \in \mathcal{K}} \|\mathbf{U}(z)\| \sup_{x \in \mathcal{X}_0} \|\mathbf{H}(\mathbf{b}^*(\alpha, x) - \mathbf{b}(\alpha, x))\| + \sup_{(x,z) \in \mathcal{X}_0 \times \mathcal{K}} |R_Q(\alpha, x, z)| \\ &= O(h^s), \end{aligned} \tag{A.8}$$

by equation (A.6) and under Assumption A.4. Define

$$A_{n,j}(\alpha, x) = \sum_{i=1}^n \{\mathbb{I}(Y_i \leq Q^*(X_i)) - \mathbb{I}(Y_i \leq Q(X_i))\} K_j\left(\frac{X_i - x}{h}\right).$$

Define $W_i = Y_i - Q(\alpha|X_i)$ and observe that (A.8) yields that

$$\begin{aligned} |A_{n,j}(\alpha, x)| &\leq nh^{s+1} \frac{1}{nh^{s+1}} \sum_{i=1}^n \mathbb{I}(|Y_i - Q(X_i)| \leq Ch^s) \left| K_j\left(\frac{X_i - x}{h}\right) \right| \\ &= nh^{s+1} \frac{1}{nh^{s+1}} \sum_{i=1}^n \mathbb{I}(|W_i| \leq Ch^s) \left| K_j\left(\frac{X_i - x}{h}\right) \right| \\ &= nh^{s+1} \tilde{f}(x, 0), \end{aligned} \tag{A.9}$$

where $\tilde{f}(x, 0)$ is a kernel estimator of the density of the couple (X_i, W_i) at $(x, 0)$. Observe that

$$\sup_{x \in \mathcal{X}_0} |\tilde{f}(x, 0)| \leq \sup_{x \in \mathcal{X}_0} |\tilde{f}(x, 0) - \mathbb{E}[\tilde{f}(x, 0)]| + \sup_{x \in \mathcal{X}_0} |\mathbb{E}[\tilde{f}(x, 0)]|. \tag{A.10}$$

We now bound the two terms in the RHS of (A.10). Einmahl and Mason ([4], Thm. 1) yields that

$$\sup_{x \in \mathcal{X}_0} |\tilde{f}(x, 0) - \mathbb{E}[\tilde{f}(x, 0)]| = O_{\mathbb{P}}\left(\frac{\log^{1/2}(n)}{(nh^{s+1})^{1/2}}\right), \tag{A.11}$$

under Assumptions A.1, A.2 and A.3. The change of variables $z = x + hu$ yields that $\mathbb{E}[\tilde{f}(x, 0)]$ can be written as

$$\int \int \mathbb{I}(|y - Q(x + hu)| \leq Ch^s) |K_j(u)| f(y|x + hu) f(x + hu) dy du.$$

It then follows from Assumptions A.1, A.2 and A.3 that $\sup_{x \in \mathcal{X}_0} |\mathbb{E}[\tilde{f}(x, 0)]| \leq C$. This, Assumption A.5 and (A.11) yield that the RHS in (A.10) is a $O_{\mathbb{P}}(1)$ which together with (A.9) proves the first part of the Lemma.

We now prove the second part of the Lemma. For that first note that

$$\sup_{x \in \mathcal{X}_0} \|\mathbf{J}_n(\alpha, x) - \mathbf{J}(\alpha, x)\| \leq \sup_{x \in \mathcal{X}_0} \|\mathbf{J}_n(\alpha, x) - \mathbb{E}[\mathbf{J}_n(\alpha, x)]\| + \sup_{x \in \mathcal{X}_0} \|\mathbb{E}[\mathbf{J}_n(\alpha, x)] - \mathbf{J}(\alpha, x)\|.$$

We control the first term in the RHS of the above equation. Observe that $\mathbf{J}_n(\alpha, x)$ is the matrix with entries

$$J_{v_1, v_2}(\alpha, x) = \sum_{i=1}^n \frac{2}{nh} K_{v_1+v_2}\left(\frac{X_i - x}{h}\right) f(Q^*(X_i)|X_i) = \sum_{i=1}^n J_{v_1, v_2, i}(\alpha, x),$$

where $0 \leq v_1, v_2 \leq p$. Using Einmahl and Mason ([4], Thm. 1) yields that

$$\sup_{x \in \mathcal{X}_0} \|J_{v_1, v_2}(\alpha, x) - \mathbb{E}[J_{v_1, v_2}(\alpha, x)]\| \leq C \left(\frac{\log(n)}{nh}\right)^{1/2},$$

for all $0 \leq v_1, v_2 \leq p$, under Assumptions A.1–A.3 and A.5. The Lemma is then proved if we show that $\sup_{x \in \mathcal{X}_0} \|\mathbb{E}[\mathbf{J}_n(\alpha, x)] - \mathbf{J}(\alpha, x)\| = O(h)$. (A.8) and the change of variables $z = x + hu$ then yield that the generic entries of $\|\mathbf{J}_n(\alpha, x) - \mathbf{J}(\alpha, x)\|$ are less than

$$\begin{aligned} & \int \left(\sup_{(x,u) \in \mathcal{X}_0 \times \mathcal{K}} |f(Q^*(x+hu)|x+hu)(f(x+hu) - f(x))| \right. \\ & \quad \left. + \sup_{(x,u) \in \mathcal{X}_0 \times \mathcal{K}} |f(x)(f(Q^*(x+hu)|x+hu) - f(Q(x)|x)f(x))| \right) |K_{v_1+v_2}(u)| du \\ & \leq C \int \left(\sup_{(x,u) \in \mathcal{X}_0 \times \mathcal{K}} |f(x+hu) - f(x)| + \sup_{(x,u) \in \mathcal{X}_0 \times \mathcal{K}} |f(Q^*(x+hu)|x+hu) - f(Q(x)|x)| \right) |K_{v_1+v_2}(u)| du \\ & = O(h), \end{aligned}$$

under Assumptions A.1, A.2 and equation (A.6).

REFERENCES

- [1] R.R. Bahadur, A note on quantiles in large samples. *Ann. Math. Stat.* **37** (1966) 577–580.
- [2] P.J. Bickel and M. Rosenblatt, On some global measures of the deviation of density function estimates. *Ann. Statist.* **1** (1973) 1071–1095.
- [3] G. Claeskens and I. Van Keilegom, Bootstrap confidence bands for regression curves and their derivatives. *Ann. Statist.* **31** (2003) 1852–1884.
- [4] U. Einmahl and D.M. Mason, Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist.* **3** (2005) 1380–1403.
- [5] R.L. Eubank and P.L. Speckman, Confidence bands in nonparametric regression. *J. Amer. Stat. Associat.* **88** (1993) 1287–1301.
- [6] J. Fan and I. Gijbels, Local Polynomial Modeling And Its Applications. *Monogr. Stat. Appl. Prob.* Chapman and Hall **66** (1996).
- [7] E. Guerre and C. Sabbah, Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. *Econom. Theory.* **28** (2012) 87–129.
- [8] W. Härdle, Asymptotic maximal deviation of M -smoothers. *J. Mult. Anal.* **29** (1989) 163–179.
- [9] W. Härdle, Y. Ritov and S. Song, Partial linear quantile regression and bootstrap confidence bands. *J. Mult. Ana.* **107** (2012) 244–262.
- [10] W. Härdle and S. Song, The Stochastic fluctuation of the quantile regression curve. *Econom. Theory* **26** (2010) 1180–1200.
- [11] P.J. Huber, Robust estimation of a location parameter. *Ann. Math. Stat.* **37** (1964) 73–101.
- [12] P.J. Huber, Robust Statistics. *Wiley Series in Probab. Math. Statist.* John Wiley and Sons, Inc., New York (1981).
- [13] G. Knaf, J. Sacks and D. Ylvisaker, Confidence bands for regression functions. *J. Amer. Stat. Associat.* **80** (1985) 683–691.
- [14] R. Koenker, *Quantile Regression*. New York, Cambridge University Press (2005).
- [15] R. Koenker and G. Basset, Regression quantiles. *Econometrica* **46** (1978) 33–50.
- [16] E. Kong, O. Linton and Y. Xia, Uniform Bahadur representation for local polynomial estimates of M -regression and its application to the additive model. *Econom. Theory.* **26** (2010) 159–166.
- [17] D.H.-Y. Leung, Cross-validation in nonparametric regression with outliers. *Ann. Statist.* **33** (2005) 2291–2310.
- [18] Q. Li and J.S. Racine, Nonparametric estimation of conditional CDF and quantile function with mixed categorical and continuous data. *J. Busin. Econ. Statist.* **26** (2008) 423–434.
- [19] R. Maronna, D. Martin and V. Yohai, *Robust statistics, theory and methods*. Wiley (2006).
- [20] J.L. Powell, Censored regression quantiles. *J. Econom.* **32** (1986) 143–155.
- [21] M. Rosenblatt, Remarks on a multivariate transformation. *Ann. Math. Stat.* **23** (1952) 470–472.
- [22] C.J. Stone, Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** (1982) 1040–1053.
- [23] J. Sun and C.R. Loader, Simultaneous confidence bands for linear regression and smoothing. *Ann. Statist.* **22** (1994) 1328–1345.
- [24] G. Tusnady, A remark on the approximation of the sample distribution function in the multidimensional case. *Period. Math. Hungar.* **8** (1977) 53–55.
- [25] J. Wang and L. Yang, Polynomial spline confidence bands for regression curves. *Statistica Sinica.* **19** (2009) 325–342.
- [26] K. Yu and M.C. Jones, Local Linear Quantile Regression. *J. Amer. Stat. Associat.* **93** (1998) 228–237.