I N S T I T U T D E S T A T I S T I Q U E

B I O S T A T I S T I Q U E E T

S C I E N C E S A C T U A R I E L L E S

(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2013/27

Expansion for Moments of Regression Quantiles with Applications to Nonparametric Testing

MAMMEN, E., VAN KEILEGOM, I. and K. YU

EXPANSION FOR MOMENTS OF REGRESSION QUANTILES WITH APPLICATIONS TO NONPARAMETRIC TESTING

Enno Mammen[§]
University of Mannheim

Ingrid Van Keilegom [**]
Université catholique de Louvain

Kyusang Yu [‡‡]
Konkuk University, Seoul

June 25, 2013

Abstract

We discuss nonparametric tests for parametric specifications of regression quantiles. The test is based on the comparison of parametric and nonparametric fits of these quantiles. The nonparametric fit is a Nadaraya-Watson quantile smoothing estimator.

An asymptotic treatment of the test statistic requires the development of new mathematical arguments. An approach that makes only use of plugging in a Bahadur expansion of the nonparametric estimator is not satisfactory. It requires too strong conditions on the dimension and the choice of the bandwidth.

Our alternative mathematical approach requires the calculation of moments of Bahadur expansions of Nadaraya-Watson quantile regression estimators. This calculation is done by inverting the problem and application of higher order Edgeworth

[§]Department of Economics, University of Mannheim, L7, 3-5, 68131 Mannheim, Germany. E-mail address: emammen@rumms.uni-mannheim.de. Tel. 0049 621 181 1927.

[**]Institute of Statistics, Biostatistics and Actuarial Sciences, Université catholique de Louvain, Voie du Roman Pays 20, B 1348 Louvain-la-Neuve, Belgium. E-mail address: ingrid.vankeilegom@uclouvain.be.

[‡‡]Department of Applied Statistics, Konkuk University, Seoul 143-701, Korea, E-mail address: kyusangu@konkuk.ac.kr

expansions. The moments allow estimation bounds for the accuracy of Bahadur expansions for integrals of kernel quantile estimators.

Another application of our method gives asymptotic results for the estimation of weighted averages of regression quantiles.

AMS 1991 subject classifications. primary 62G07, secondary 62G20

Journal of Economic Literature Classification: C14

Keywords and phrases. Nonparametric Regression; Quantiles; Kernel Smoothing; Nonparametric Testing; Goodness-of-fit tests.

1 Introduction

Consider a data set of n i.i.d. tuples (X_i, Y_i), where Y_i is a one-dimensional response variable and X_i is a d-dimensional covariate. For $0 < \alpha < 1$ we denote the conditional α-quantile of Y_i given $X_i = x$ by $m_\alpha(x)$. Thus we can write

$$Y_i = m_\alpha(X_i) + \varepsilon_{i,\alpha} \quad (i = 1, \ldots, n), \tag{1}$$

with error variables $\varepsilon_{i,\alpha}$ that fullfill $q_\alpha(\varepsilon_{i,\alpha}|X_i) = 0$. Here, $q_\alpha(\varepsilon_{i,\alpha}|X_i)$ is the α-quantile of the conditional distribution of $\varepsilon_{i,\alpha}$ given X_i. Consider the null hypothesis

$$H_0 : \text{For all } \alpha \in A \text{ there exists a } \theta(\alpha) \in \Theta, \text{ such that : } m_\alpha = m_{\alpha,\theta(\alpha)}, \tag{2}$$

where $\{m_{\alpha,\theta} : \theta \in \Theta\}$ is a parametric class of regression quantiles, Θ is a compact subset of $\mathbb{R}^k$ and $A \subset (0, 1)$. The set A can be a singleton $A = \{\alpha\}$, but can also be a (closed) subset of $(0, 1)$ if a set of quantile functions is checked.

In this paper we aim at studying a test statistic for H_0, and to study its asymptotic properties under the null and the alternative. We will see that this problem is an example of a quantile model where the asymptotics cannot be developed by standard tools of quantile regression. In particular, a direct application of Bahadur expansions requires assumptions that are too restrictive.

Our test is an omnibus test that has power against all types of alternatives. It is based on the comparison of a kernel quantile estimator with the parametric fit. The test statistic is a weighted L_2-distance between the nonparametric and the parametric estimator. Similar tests have been used in a series of papers for mean regression. Early references are Härdle and Mammen (1993), González-Manteiga and Cao-Abad (1993), Hjellvik, Yao and Tjøstheim (1998), Zheng (1996) and Fan, Zhang and Zhang (2001). Furthermore recent references are Dette and Sprekelsen (2004), Kreiss, Neumann and Yao (2008), Haag (2008), Leucht (2012), Gao and Hong (2008) and Ait-Sahalia, Fan and Peng (2009). Most of the more recent work concentrates on time series data.

The classical way to carry over results from parametric and nonparametric mean regression to quantile regression is the use of Bahadur expansions. The main point is that asymptotically quantile regression is equivalent to weighted mean regression. This approach has been used in Chaudhuri (1991), Truong (1989), He and Ng (1999), He, Ng and Portnoy (1998) and more recently in Hoderlein and Mammen (2009), Hong (2003), Kong, Linton and Xia (2010), Lee and Lee (2008), El Ghouch and Van Keilegom (2009) and Li and Racine (2008). A detailed review of quantile regression can be found in the book by Koenker (2005). Testing procedures in quantile regression were considered in He and Zhu (2003), Koenker and Machado (1999), Koenker and Xiao (2002) and Zheng (1998).

In this paper we will discuss how results from mean regression carry over to our case. Whereas elsewhere a first attempt could be based on the application of a Bahadur expansion, we will see that in our setting the accuracy of a direct application of Bahadur expansions is too poor. Such an approach would require that the bandwidth h of the nonparametric kernel regression quantile estimator fulfills that $nh^{3d} \to \infty$ for sample size n going to ∞. Here, d is the dimension of the covariate. E.g. if one applies a bandwidth $h \sim n^{-1/(4+d)}$ that leads to a rate optimal estimation of twice differentiable functions this assumption would allow only a one-dimensional setting $d = 1$. Also in the case of minimax optimal testing with twice differentiable functions under the alternative (see

Ingster (1993) and Guerre and Lavergne (2002)), the optimal bandwidth $h \sim n^{-2/(8+d)}$ is only allowed for dimension $d = 1$. In this paper we develop an asymptotic theory for L_2-type quantile tests that works under the assumption that $nh^{3d/2} \to \infty$. In the above examples this allows dimensions $d \leq 7$ and $d \leq 3$. Furthermore, we will argue that the assumption $nh^{3d/2} \to \infty$ is necessary to get the same asymptotics for L_2-type quantile tests as for L_2-type mean regression tests. We will shortly outline that our theory can be used to go beyond the assumption $nh^{3d/2} \to \infty$ and how the asymptotics of the test statistics would change in this case. In our approach we will make use of the fact that Bahadur expansions of kernel quantile estimators are asymptotically independent if they are calculated at points that differ more than twice the bandwidth h. Thus the variance of an integral over a Bahadur expansion should be of smaller order than the variance of the Bahadur expansion at a fixed point. The main technical difficulty that will come up when applying this idea is the need to calculate moments of the Bahadur expansion. We will introduce a method for the expansions of such moments that is based on Edgeworth expansions in a related problem.

The paper is organized as follows. In the next section we will state our main result on the asymptotics of L_2-type quantile tests. We will also introduce some kind of wild bootstrap procedure adapted to quantile regression and give a theoretical result on its consistency. Our theory only applies for Nadaraya-Watson type smoothing. We add a result on tests based on local polynomial smoothing. This result is based on a direct application of Bahadur expansions and requires the stronger condition $nh^{3d} \to \infty$. The proofs are postponed to the last three sections.

2 Asymptotic theory

Recall that we are interested in the null hypothesis H_0 defined in (2). We suppose that for all $\alpha \in A$,

$$m_\alpha(\cdot) = m_{\alpha,\theta_0(\alpha)}(\cdot) + n^{-1/2}h^{-d/4}\Delta_\alpha(\cdot). \tag{3}$$

For the case $\Delta_\alpha \equiv 0$ the function m_α lies on the hypothesis. We suppose that there exists an estimator $\widehat{\theta}(\alpha)$ that converges to $\theta_0(\alpha)$. Hence, on the hypothesis the true value of $\theta(\alpha)$ is equal to $\theta_0(\alpha)$. On the alternative, $\theta_0(\alpha)$ may depend on the chosen estimator $\widehat{\theta}(\alpha)$.

Let $K(u_1, \ldots, u_d) = \prod_{j=1}^d k(u_j)$, where k is a one-dimensional density function defined on $[-1, 1]$, and let $h = (h_1, \ldots, h_d)$ be a d-dimensional bandwidth parameter. We assume that all bandwidths $h_1, \ldots, h_d$ are of the same order. For simplicity of notation we further assume that they are identical and by abuse of notation we write $h = h_1 = \ldots = h_d$. For any $0 < \alpha < 1$ and any x in the support R_X of X, let $F_{\varepsilon_\alpha|X}(\cdot|x)$ be the conditional distribution function of $\varepsilon_\alpha = Y - m_\alpha(X)$, given $X = x$, and let $r_{\alpha,\theta(\alpha)}(x)$ be the α-quantile of $Y - m_{\alpha,\theta(\alpha)}(X)$ given that $X = x$. Define

$$\widehat{r}_\alpha(x) = \arg\min_r \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \tau_\alpha(Y_i - m_{\alpha,\widehat{\theta}(\alpha)}(X_i) - r),$$

where $\tau_\alpha(u) = \alpha u_+ - (1-\alpha)u_-$, $u_+ = uI(u > 0)$ and $u_- = uI(u < 0)$.

We suppose that A is a closed subinterval of $(0, 1)$. We define the following test statistic :

$$\widehat{T}_A = \int_A \int_{R_X} \widehat{r}_\alpha(x)^2 w(x, \alpha) dx d\alpha, \tag{4}$$

for some weight function $w(x, \alpha)$. For the case that A contains only one value α we use

$$\widehat{T}_\alpha = \int_{R_X} \widehat{r}_\alpha(x)^2 w(x) dx \tag{5}$$

for some weight function $w(x)$. One could also generalize our results to the case that A is a finite set. To keep notation simple we omit this case in our mathematical analysis.

In order to develop the asymptotic distribution of $\widehat{T}_A$ and $\widehat{T}_\alpha$, we need to work under the following assumptions. In the formulation of the assumptions and in the proofs we use the convention that $C, C_1, C_2, \ldots$ are generic strictly positive constants that are chosen large enough, that $c, c_1, c_2, \ldots$ are generic strictly positive constants that are chosen small enough, and that $C^*, C_1^*, C_2^*, \ldots$ are generic strictly positive constants that are arbitrarily chosen. Using this convention we write $L_n = (\log n)^C$ for a sequence with $C > 0$ large enough and $L_n^* = (\log n)^{C^*}$ for a sequence with an arbitrarily chosen constant $C^* > 0$. All these variable names are used for different constants and sequences, even in the same equation.

(B1) The support R_X of X is a compact convex subset of $I\!R^d$. The density f_X of X is strictly positive and continuously differentiable on the interior of R_X. The conditional density $f_{X|\varepsilon_\alpha + n^{-1/2}h^{-d/4}\Delta_\alpha(X)}(x|e)$ of X given $\varepsilon_\alpha + n^{-1/2}h^{-d/4}\Delta_\alpha(X) = e$ is uniformly bounded over x, e, n and α for n large enough.

(B2) The cumulative distribution function $F(\cdot|x)$ of the conditional distribution of Y given $X = x$ is continuously differentiable with respect to x and has a density $f(\cdot|x)$ that satisfies

$$\begin{aligned} f(y|x) &> 0, \\ |f(y'|x') - f(y|x)| &\leq C(\|x' - x\| + |y' - y|) \end{aligned}$$

for $x, x' \in R_X$ and $y, y' \in \mathbb{R}$, where $\|\cdot\|$ is the Euclidean norm.

(B3) The kernel k is a symmetric, continuously differentiable probability density function with compact support, $[-1, 1]$, say. It fulfills a Lipschitz condition and it is monotone strictly increasing on $[-1, 0]$. It holds that $k'(k^{-1}(u)) \geq \min c\{u^\kappa, (k(0) - u)^\kappa\}$ for some $\kappa > 0$ where $k^{-1} : [0, k(0)] \to [-1, 0]$ denotes the inverse of $k : [-1, 0] \to [0, k(0)]$. The bandwidth h satisfies $h = o(1)$ and $nh^{3d/2}/L_n \to \infty$.

(B4) We assume that

$$\sup_{x \in R_X, \alpha \in A} |m_{\alpha,\widehat{\theta}(\alpha)}(x) - m_{\alpha,\theta_0(\alpha)}(x) - (\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top \gamma_\alpha(x)| = O_P(n^{-\frac{1}{2}-c})$$

for some function $\gamma_\alpha(x)$. The functions $w(x, \alpha)$, $\gamma_\alpha(x)$ and $\Delta_\alpha(x)$ are continuous with respect to α and x. The function $w(x)$ is continuous. For $g(x) = w(x, \alpha)$ and $g(x) = w(x)$ it holds that $|g(x') - g(x)| \leq C\|x' - x\|$, and $|\gamma_\alpha(x') - \gamma_\alpha(x)| \leq C\|x' - x\|^\delta$ for some $0 < \delta < 1$ and for all $x, x' \in R_X$ and all $\alpha \in A$.

(B5) It holds that $\sup_{\alpha \in A} \|\widehat{\theta}(\alpha) - \theta_0(\alpha)\| = O_P(n^{-\frac{1}{2}+c})$. For $g(\alpha) = \theta_0(\alpha)$ it holds that $\|g(\alpha') - g(\alpha)\| \leq C|\alpha' - \alpha|$ for $\alpha, \alpha' \in A$. For $g(\alpha) = \widehat{\theta}(\alpha)$ this holds with probability tending to one.

We can now state our first main result.

Theorem 1. *Assume (B1)-(B5). Then,*

$$nh^{d/2}\widehat{T}_A - b_{h,A} \xrightarrow{d} N(D_A, V_A),$$
$$nh^{d/2}\widehat{T}_\alpha - b_{h,\alpha} \xrightarrow{d} N(D_\alpha, V_\alpha),$$

where

$$\begin{aligned}
D_A &= \int_A \int_{R_X} \Delta_\alpha(x)^2 w(x,\alpha)\, dx\, d\alpha, \\
b_{h,A} &= h^{-d/2} K^{(2)}(0) \int_A \alpha(1-\alpha) \int_{R_X} \frac{w(x,\alpha)}{f_X(x) f^2_{\varepsilon_\alpha|X}(0|x)} dx\, d\alpha, \\
V_A &= 4K^{(4)}(0) \int_{\alpha,\beta \in A, \alpha < \beta} \alpha^2 (1-\beta)^2 \int_{R_X} \frac{w(x,\alpha) w(x,\beta)}{f^2_X(x) f^4_{\varepsilon_\alpha|X}(0|x)} dx\, d\alpha\, d\beta, \\
D_\alpha &= \int_{R_X} \Delta_\alpha(x)^2 w(x)\, dx, \\
b_{h,\alpha} &= h^{-d/2} K^{(2)}(0) \alpha(1-\alpha) \int_{R_X} \frac{w(x)}{f_X(x) f^2_{\varepsilon_\alpha|X}(0|x)} dx, \\
V_\alpha &= 4K^{(4)}(0) \alpha^2 (1-\alpha)^2 \int_{R_X} \frac{w^2(x)}{f^2_X(x) f^4_{\varepsilon_\alpha|X}(0|x)} dx,
\end{aligned}$$

and where for any j, $K^{(j)}(0)$ denotes the j-times convolution product of K at 0.

In our theorem we make the assumption that $nh^{3d/2}/L_n$ converges to ∞. This assumption is used in the proof of Lemma 5 in Section 3. For the calculation of the conditional second moment of $\widehat{r}$ we used an Edgeworth expansion that gives an expansion for the moment with remainder term $O_P(n^{-2}h^{-2d})$. This term is of order $o_P(n^{-1}h^{-d/2})$ if $nh^{3d/2} \to \infty$. This suffices for the asymptotic result of Theorem 1. If it does not hold that $nh^{3d/2} \to \infty$ we get an additional bias term for $\widehat{T}_A$ and $\widehat{T}_\alpha$ that is of order $O(n^{-2}h^{-2d})$. This would require a higher order expansion in Lemma 5. In particular, we conjecture that we then get an asymptotic limit for our test statistic that differs from the limit if the estimator $\widehat{r}_\alpha$ is replaced by its Bahadur expansion.

We expect that Theorem 1 cannot be used for an accurate calculation of critical values. The asymptotic normality result of Theorem 1 is based on the fact that kernel smoothers are asymptotically independent if they are calculated at points that differ more than $2h$. Thus the convergence is comparable to the convergence of the sum of h^{-d} independent summands. This would motivate a rate of convergence of order $h^{-d/2}$. As has been suggested for other goodness-of-fit tests in the literature, also here a way out is to use a bootstrap procedure. We will introduce some kind of wild bootstrap for quantiles in which the Bahadur expansion $\widetilde{r}_\alpha$ of $\widehat{r}_\alpha$ is resampled. For the definition of $\widetilde{r}_\alpha$ see (7) in Section 3. For the bootstrap, we define

$$\widetilde{r}^*_\alpha(x) = -\frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)\{I(U_i \le \alpha) - \alpha\}}{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) f^*_{\varepsilon^*_\alpha|X^*}(0|X_i)},$$

where $f^*_{\varepsilon^*_\alpha|X^*}$ is an estimator of $f_{\varepsilon_\alpha|X}$ and U_i are independent random variables with uniform distribution on $[0,1]$ that are independent of the sample. The bootstrap test statistics are defined as:

$$\widehat{T}^*_A = \int_A \int_{R_X} \widetilde{r}^*_\alpha(x)^2 w(x,\alpha) dx \; d\alpha$$

and

$$\widehat{T}^*_\alpha = \int_{R_X} \widetilde{r}^*_\alpha(x)^2 w(x) dx.$$

For proving the consistency of this bootstrap procedure, we need one more assumption :

(B6) It holds that

$$\sup_{\alpha \in A,\; x \in R_X} \left| f^*_{\varepsilon^*_\alpha|X^*}(0|x) - f_{\varepsilon_\alpha|X}(0|x) \right| \to 0,$$

in probability.

The next theorem shows the consistency of the above bootstrap approach.

Theorem 2. *Assume (B1)-(B6). Then,*

$$d_K(\mathcal{L}^*(nh^{d/2}\widehat{T}_A^* - b_{h,A}), N(0, V_A)) \xrightarrow{p} 0,$$
$$d_K(\mathcal{L}^*(nh^{d/2}\widehat{T}_\alpha^* - b_{h,\alpha}), N(0, V_\alpha)) \xrightarrow{p} 0,$$

where $\mathcal{L}^*(...)$ denotes the conditional distribution, given the sample. Furthermore, d_K is the Kolmogorov distance, i.e. the sup norm of the difference between the corresponding distribution functions.

Theorem 2 remains to hold if we replace (B1)–(B5) by weaker conditions. We do not pursuit this because we need for consistency of bootstrap that both, Theorem 1 and Theorem 2, hold.

Our test proposed above is based on local constant smoothing and it checks for the accuracy of parametric specifications of the regression function. We now shortly discuss tests based on local polynomial smoothing of order p. We are doing this for two reasons. First of all it allows for checking the accuracy of estimated derivatives of the regression function. Second, since this result is based on a direct application of Bahadur expansions, we will see that this leads to much more restrictive assumptions on the bandwidth than we had in Theorem 1. We remark that the approach of Theorem 1 cannot be applied for local polynomials. The reason is that an equation like (11) in Section 3 does not hold for local polynomials because the local polynomial method fits a local vector instead of a local scalar.

Define for d-dimensional vectors z and ν,

$$\pi(z)^\top b = \sum_{\nu:|\nu|\le p} b_\nu \frac{z^\nu}{\nu!},$$
$$\pi_h(z)^\top b = \sum_{\nu:|\nu|\le p} b_\nu \frac{z^\nu}{h^{|\nu|}\nu!},$$

with the convention $z^\nu = \prod_{j=1}^d z_j^{\nu_j}$, $|\nu| = \sum_{j=1}^d \nu_j$ and $\nu! = \prod_{j=1}^d \nu_j!$. We let r_p be the number of elements of $R_p = \{\nu : |\nu| \le p\}$ and we define an r_p-dimensional vector $e_{\nu^\dagger}$ that has elements 0 for $\nu \ne \nu^\dagger$ and an entry equal to 1 for $\nu = \nu^\dagger$. Here the elements of the vector are ordered according to some listing of the set R_p. Now, the test statistic is based on the following local polynomial residual quantile estimator:

$$\widehat{r}_\alpha(x) = \arg\min_b \sum_{i=1}^n K\Big(\frac{x - X_i}{h}\Big)\tau_\alpha(Y_i - m_{\alpha,\widehat{\theta}(\alpha)}(X_i) - \pi_h(X_i - x)^\top b).$$

Here $\widehat{r}_\alpha(x)$ is an r_p-dimensional vector and for fixed $\nu^\dagger$, we put

$$\widehat{r}^\dagger_\alpha(x) = \widehat{r}_{\alpha,\nu^\dagger}(x).$$

In analogy to the above definitions, the test for the parametric model is based on one of the following two test statistics: $\widehat{T}^\dagger_A = \int_A \int_{R_X} \widehat{r}^\dagger_\alpha(x)^2 w(x,\alpha)dx\ d\alpha$ or $\widehat{T}^\dagger_\alpha = \int_{R_X} \widehat{r}^\dagger_\alpha(x)^2 w(x)dx$, respectively. Note that $h^{-|\nu^\dagger|}\widehat{r}^\dagger_\alpha(x)$ is an estimator of the $\nu^\dagger$-th partial derivative of $m_\alpha(x) - m_{\alpha,\theta_0(\alpha)}(x)$. Thus the test checks for the accuracy of the parametric fit of the $\nu^\dagger$-th partial derivative of m_α. Instead of (3) we now assume that

$$m_\alpha(\cdot) = m_{\alpha,\theta_0(\alpha)}(\cdot) + n^{-1/2}h^{-d/4-|\nu^\dagger|}\Delta_\alpha(\cdot). \tag{6}$$

Note that for $|\nu^\dagger| > 0$ we get slower rates of convergence for alternatives that are detected with non trivial power. We conjecture that this difference in power rates disappears if one considers other classes of alternatives. In particular, for the alternative $m_\alpha(\cdot) = m_{\alpha,\theta_0(\alpha)}(\cdot) + c_n\Delta_\alpha(\cdot/h)$ we expect that the test has nontrivial power for the same rate c_n for all values of $|\nu^\dagger|$.

We need the following assumptions for the asymptotic results of our next theorem.

(B7) The kernel k is a probability density function with compact support. It fulfills a Lipschitz condition and it holds that $k(0) > 0$. For the bandwidth h we have that $h = h_n \to 0$ and $nh^{3d}/L_n \to \infty$.

(B8) $\Delta_\alpha(x)$ has a derivative $\Delta_\alpha^{(\nu^\dagger)}(x)$ with respect to x of order $\nu^\dagger$ that is continuous in α and x.

We are now ready to state an asymptotic result regarding the tests based on $\widehat{T}^\dagger_A$ and $\widehat{T}^\dagger_\alpha$, respectively.

Theorem 3. *Assume (B1), (B2), (B4) (B5), (B7), (B8). Then,*

$$nh^{d/2+2|\nu^\dagger|}\widehat{T}_A^\dagger - b_{h,A}^\dagger \xrightarrow{d} N(D_A^\dagger, V_A^\dagger),$$
$$nh^{d/2+2|\nu^\dagger|}\widehat{T}_\alpha^\dagger - b_{h,\alpha}^\dagger \xrightarrow{d} N(D_\alpha^\dagger, V_\alpha^\dagger),$$

where

$$\begin{aligned}
D_A^\dagger &= \int_A \int_{R_X} \Delta_\alpha^{(\nu^\dagger)}(x)^2 w(x,\alpha)\,dx\,d\alpha,\\
b_{h,A}^\dagger &= h^{-d/2}L^{(2)}(0)\int_A \alpha(1-\alpha)\int_{R_X}\frac{w(x,\alpha)}{f_X(x)f^2_{\varepsilon_\alpha|X}(0|x)}dx\,d\alpha,\\
V_A^\dagger &= 4L^{(4)}(0)\int_{\alpha,\beta\in A,\alpha<\beta}\alpha^2(1-\beta)^2\int_{R_X}\frac{w(x,\alpha)w(x,\beta)}{f_X^2(x)f^4_{\varepsilon_\alpha|X}(0|x)}dx\,d\alpha\;d\beta,\\
D_\alpha^\dagger &= \int_{R_X}\Delta_\alpha^{(\nu^\dagger)}(x)^2 w(x)\,dx,\\
b_{h,\alpha}^\dagger &= h^{-d/2}L^{(2)}(0)\alpha(1-\alpha)\int_{R_X}\frac{w(x)}{f_X(x)f^2_{\varepsilon_\alpha|X}(0|x)}dx,\\
V_\alpha^\dagger &= 4L^{(4)}(0)\alpha^2(1-\alpha)^2\int_{R_X}\frac{w^2(x)}{f_X^2(x)f^4_{\varepsilon_\alpha|X}(0|x)}dx,
\end{aligned}$$

where the kernel L is defined as $L(u) = e_{\nu^\dagger}^\top\left[\int \pi(v)\pi(v)^\top K(v)\;dv\right]^{-1}\pi(u)K(u)$.

3 Proof of Theorem 1

We only prove the statement for $\widehat{T}_A$. The asymptotic result for $\widehat{T}_\alpha$ follows similarly. We need to introduce a few more notations. With $\delta_{\theta,\alpha}(x) = -(\theta(\alpha)-\theta_0(\alpha))^\top\gamma_\alpha(x) + n^{-1/2}h^{-d/4}\Delta_\alpha(x)$ and $\varepsilon_{i,\alpha}^\Delta = \varepsilon_{i,\alpha} + n^{-1/2}h^{-d/4}\Delta_\alpha(X_i)$ we put

$$\begin{aligned}
\widetilde{r}_\alpha(x) &= -\frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)\{I(\varepsilon_{i,\alpha}^\Delta \le 0)-\alpha\}}{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)f_{\varepsilon_\alpha|X}(0|X_i)}, \qquad (7)\\
\widehat{r}_{\alpha,\theta}(x) &= \arg\min_r \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)\tau_\alpha(\varepsilon_{i.\alpha}+\delta_{\theta,\alpha}(X_i)-r),\\
\widehat{r}_\alpha^0(x) &= \widehat{r}_{\alpha,\theta_0}(x)
\end{aligned}$$

and

$$\widehat{r}_\alpha^\star(x) = \begin{cases} \widehat{r}_\alpha^0(x) & \text{if } |\widehat{r}_\alpha^0(x)| \le L_n(nh^d)^{-1/2}\\ 0 & \text{otherwise.}\end{cases}$$

Let also

$$W_{ni}(x,h) \quad = \quad K_h(x-X_i)/[\sum_j K_h(x-X_j)],$$

with $K_h(\cdot) = K(\cdot/h)/h^d$.

If X would be one-dimensional, the support R_X would be a compact interval. For arbitrary j and for $k = 1, 2, 3$, we can then define

$$I_{jk} = [(3j+k-1)h, (3j+k)h], \qquad \text{and} \qquad I^*_{jk} = [(3j+k-2)h, (3j+k+1)h].$$

The set of indices of the X_i $(i = 1, \ldots, n)$ that fall inside the interval I^*_{jk} is denoted by $\mathcal{N}_{jk}$. We write N_{jk} for the number of elements of $\mathcal{N}_{jk}$. An arbitrary $x \in R_X$ belongs to a unique I_{jk} and we define $\mathcal{N}(x) = \mathcal{N}_{jk}$ and $N(x) = N_{jk}$. If the dimension of X is larger than one, this partition of the support into small intervals can be generalized in an obvious way.

The proof of Theorem 1 will make use of the following lemmas.

Lemma 1. *Suppose that the assumptions of Theorem 1 are satisfied. Then,*

$$\sup_{\alpha\in A}\sup_{x\in R_X}\left|\widehat{r}_{\alpha,\widehat{\theta}(\alpha)}(x)\right| = O_P((nh^d)^{-1/2}L_n),$$

$$\sup_{\alpha\in A}\sup_{x\in R_X}\left|\widehat{r}^0_{\alpha}(x)\right| = O_P((nh^d)^{-1/2}L_n).$$

Proof of Lemma 1. As is known for the case where there is no parametric part and where $\Delta_\alpha \equiv 0$, one has that

$$\sup_{\alpha\in A}\sup_{x\in R_X}\left|\widehat{r}_{\alpha}(x) - \widetilde{r}_{\alpha}(x)\right| = O_P((nh^d)^{-3/4}L_n).$$

For a proof see Theorem 2 in Guerre and Sabbah (2012). By standard smoothing theory we have that (still when $\Delta_\alpha \equiv 0$)

$$\sup_{\alpha\in A}\sup_{x\in R_X}\left|\widetilde{r}_{\alpha}(x)\right| = O_P((nh^d)^{-1/2}L_n). \tag{8}$$

We can move from this case to $\widehat{r}^0_{\alpha}(x)$ by adding to the observations terms of order $O_P(n^{-1/2}h^{-d/4}) = O_P((nh^d)^{-1/2}L_n)$. In the case of $\widehat{r}_{\alpha,\widehat{\theta}(\alpha)}(x)$ we have to add to the observations terms of the order $O_P(n^{-\frac{1}{2}+c}) + O_P(n^{-1/2}h^{-d/4}) = O_P((nh^d)^{-1/2}L_n)$. This changes the local quantiles by at most this amount. This shows the statements of the lemma. □

Lemma 2. *Suppose that the assumptions of Theorem 1 are satisfied. Then,*

$$\sup_{\alpha\in A}\sup_{x\in R_X}\left|\widehat{r}_\alpha(x) - \widehat{r}^0_\alpha(x) + (\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top\gamma_\alpha(x)\right| = O_P(n^{-\frac{1}{2}-c}).$$

Proof of Lemma 2. First note that $\widehat{r}_\alpha(x)+(\widehat{\theta}(\alpha)-\theta_0(\alpha))^\top\gamma_\alpha(x)$ is equal to the quantile estimator we would obtain when we shift all observations Y_i in the window around x by the amount $(\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top\gamma_\alpha(x)$, and hence we need to show that the distance between this latter estimator (say $\widehat{r}_{\alpha,mod}(x)$) and $\widehat{r}^0_\alpha(x)$ is $O_P(n^{-\frac{1}{2}-c})$ uniformly in α and x.

Next, note that if now in addition we perturb all observations in the window around x by adding $m_{\alpha,\widehat{\theta}(\alpha)}(X_i) - m_{\alpha,\theta_0(\alpha)}(X_i) - (\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top\gamma_\alpha(X_i)$, the quantile estimator $\widehat{r}_{\alpha,mod}(x)$ will get perturbed by at most the maximal perturbation of the observations, which is of the order $O_P(n^{-1/2-c})$ by Assumption (B4).

After these two perturbations, the quantile estimator is now based on $Y_i-m_{\alpha,\theta_0(\alpha)}(X_i)+(\widehat{\theta}(\alpha)-\theta_0(\alpha))^\top(\gamma_\alpha(x)-\gamma_\alpha(X_i))$ instead of $Y_i - m_{\alpha,\widehat{\theta}(\alpha)}(X_i)$. Finally note that if we apply one more perturbation by subtracting $(\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top(\gamma_\alpha(x) - \gamma_\alpha(X_i))$ for all X_i in the window around x, the estimator changes by at most $O_P(n^{-1/2+c}h^\delta)$ by Assumption (B5), and this is $O_P(n^{-1/2-c})$ for c small enough. The so-obtained estimator equals $\widehat{r}^0_\alpha(x)$, which shows the statement of the lemma. □

Lemma 3. *Suppose that the assumptions of Theorem 1 are satisfied. Then,*

$$\sup_{\alpha\in A}\sup_{x\in R_X}\left|\widehat{r}^0_\alpha(x) - \widetilde{r}_\alpha(x)\right| = O_P((nh^d)^{-3/4}L_n).$$

Proof of Lemma 3. Write

$$\begin{aligned}
&|\widehat{r}^0_\alpha(x) - \widetilde{r}_\alpha(x)| \\
&\le \frac{1}{\inf_{x,\alpha} f_{\varepsilon_\alpha|X}(0|x)}\Big|\sum_{i=1}^n W_{ni}(x,h)f_{\varepsilon_\alpha|X}(0|X_i)\widehat{r}^0_\alpha(x) + \sum_{i=1}^n W_{ni}(x,h)\big(I(\varepsilon^\Delta_{i,\alpha}\le 0) - \alpha\big)\Big| \\
&= \frac{1}{\inf_{x,\alpha} f_{\varepsilon_\alpha|X}(0|x)}\Big|\sum_{i=1}^n W_{ni}(x,h)f_{\varepsilon_\alpha|X}(0|X_i)\widehat{r}^0_\alpha(x) - \widehat{F}_{\varepsilon^\Delta_\alpha|X}(\widehat{r}^0_\alpha(x)|x) + \widehat{F}_{\varepsilon^\Delta_\alpha|X}(0|x)\Big| \\
&\quad + O_P((nh^d)^{-1}), \qquad (9)
\end{aligned}$$

where $\widehat{F}_{\varepsilon^\Delta_\alpha|X}(y|x) = \sum_i W_{ni}(x,h) I(\varepsilon^\Delta_{i,\alpha} \leq y)$. The latter equality follows from the fact that

$$|\widehat{F}_{\varepsilon^\Delta_\alpha|X}(\widehat{r}^0_\alpha(x)|x) - \alpha| \leq |\widehat{F}_{\varepsilon^\Delta_\alpha|X}(\widehat{r}^0_\alpha(x)|x) - \widehat{F}_{\varepsilon^\Delta_\alpha|X}(\widehat{r}^0_\alpha(x) - |x)| = O_P((nh^d)^{-1}).$$

The following expansion follows from standard kernel smoothing theory, uniformly for $x \in R_X, \alpha \in A, |y| \leq a_n$ and for sequences a_n with $a_n^{-1} = O(nh^d)$:

$$\begin{aligned}
&\widehat{F}_{\varepsilon^\Delta_\alpha|X}(y|x) - \widehat{F}_{\varepsilon^\Delta_\alpha|X}(0|x) \\
&= \sum_i W_{ni}(x,h) \int_0^y f_{\varepsilon_\alpha|X}(u - n^{-1/2}h^{-d/4}\Delta_\alpha(X_i)|X_i)du + O_P((nh^d)^{-1/2}L_n a_n^{1/2}) \\
&= \sum_i W_{ni}(x,h) \int_0^y f_{\varepsilon_\alpha|X}(u|X_i)du + O_P((nh^d)^{-1/2}L_n a_n^{1/2}) + O_P(n^{-1/2}h^{-d/4}a_n) \\
&= y \sum_i W_{ni}(x,h) f_{\varepsilon_\alpha|X}(0|X_i) + O_P((nh^d)^{-1/2}L_n a_n^{1/2} + a_n^2) + O_P(n^{-1/2}h^{-d/4}a_n).
\end{aligned}$$

We now apply this bound to $a_n = (nh^d)^{-1/2}L_n$ and $y = \widehat{r}^0_\alpha(x)$, which is possible thanks to Lemma 1. This combined with (9) shows the statement of the lemma. □

For proving Theorem 1, we will make use of the following decomposition, which follows from Lemma 2 :

$$\begin{aligned}
\widehat{T}_A &= \int_A \int_{R_X} \Big[\widehat{r}^0_\alpha(x) - (\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top \gamma_\alpha(x)\Big]^2 w(x,\alpha)\,dx\,d\alpha + o_P(n^{-1}h^{-d/2}) \\
&= \int_A \int_{R_X} \Big[\widehat{r}^0_\alpha(x)^2 - \widehat{r}^\star_\alpha(x)^2\Big] w(x,\alpha)\,dx\,d\alpha \\
&\quad + \int_A \int_{R_X} E\Big\{\widehat{r}^\star_\alpha(x)^2 - \widetilde{r}_\alpha(x)^2 \Big| N(x)\Big\} w(x,\alpha)\,dx\,d\alpha \\
&\quad + \int_A \int_{R_X} \Big[\widehat{r}^\star_\alpha(x)^2 - \widetilde{r}_\alpha(x)^2 - E\Big\{\widehat{r}^\star_\alpha(x)^2 - \widetilde{r}_\alpha(x)^2 \Big| N(x)\Big\}\Big] w(x,\alpha)\,dx\,d\alpha \\
&\quad -2 \int_A \int_{R_X} \Big[(\widehat{r}^0_\alpha(x) - \widetilde{r}_\alpha(x))\big\{(\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top \gamma_\alpha(x)\big\}\Big] w(x,\alpha)\,dx\,d\alpha \\
&\quad -2 \int_A \int_{R_X} \Big[\widetilde{r}_\alpha(x)\big\{(\widehat{\theta}(\alpha) - \theta_0(\alpha))^\top \gamma_\alpha(x)\big\}\Big] w(x,\alpha)\,dx\,d\alpha
\end{aligned}$$

$$+\int_A\int_{R_X}\Big[(\widehat{\theta}(\alpha)-\theta_0(\alpha))^\top\gamma_\alpha(x)\Big]^2 w(x,\alpha)\,dx\,d\alpha$$
$$+\int_A\int_{R_X}\widetilde{r}_\alpha(x)^2 w(x,\alpha)\,dx\,d\alpha+o_P(n^{-1}h^{-d/2})$$
$$=T_{n1}+...+T_{n7}+o_P(n^{-1}h^{-d/2}),$$

where for any ℓ we denote by $E(\cdot|N(x)=\ell)$ the expected value given that the interval I_{jk} to which x belongs contains $N(x)=\ell$ elements.

Lemma 4. *Suppose that the assumptions of Theorem 1 are satisfied. Then,*

$$T_{n1}=o_P(a_n),$$

for any sequence $\{a_n\}$ of positive constants tending to zero as $n\to\infty$.

Proof of Lemma 4. Note that

$$T_{n1}\le\sup_{\alpha\in A}\sup_{x\in R_X}|\widehat{r}^0_\alpha(x)|^2\int_A\int_{R_X}I\Big(|\widehat{r}^0_\alpha(x)|>L_n(nh^d)^{-1/2}\Big)w(x,\alpha)\,dx\,d\alpha.$$

It is easily seen from Lemma 1 that

$$\int_A\int_{R_X}I\Big(|\widehat{r}^0_\alpha(x)|>L_n(nh^d)^{-1/2}\Big)w(x,\alpha)\,dx\,d\alpha=o_P(a_n),$$

for any $a_n\to 0$, since the indicator inside the integral will be zero from some point on. □

Lemma 5. *Suppose the assumptions of Theorem 1 are satisfied. Then,*

$$\sup_{\alpha\in A}\sup_{x\in R_X}\left|E\Big\{\widehat{r}^\star_\alpha(x)^2-\widetilde{r}_\alpha(x)^2\Big|N(x)\Big\}\right|=o_P((nh^{d/2})^{-1}),$$

and hence, $T_{n2}=o_P((nh^{d/2})^{-1})$.

Proof of Lemma 5. Put $\mathcal{N}^-(x)=\{u:x_j-h\le u_j\le x_j+h \text{ for all } j=1,\ldots,d\}$. This is the support of the kernel $h^{-d}K(h^{-1}[x-\cdot])$. We also write $N^-(x)$ for the random number of X_i's that lie in $\mathcal{N}^-(x)$. Note that $\mathcal{N}^-(x)\subset\mathcal{N}(x)$ and $N^-(x)\le N(x)$. We use the shorthand notation $m_0=nh^d$.

We will show below that

$$E\Big\{\widehat{r}^\star_\alpha(x)^2-\widetilde{r}_\alpha(x)^2\Big|N^-(x)=m\Big\}=O(L_n n^{-3/2}h^{-5d/4}),\qquad(10)$$

uniformly in $x \in R_X$, $\alpha \in A$ and $C_1^* m_0 \le m \le C_2^* m_0$. For $m^+ \ge m$ we have by a simple argument that $E\Big\{\widehat{r}_\alpha^\star(x)^2 - \widetilde{r}_\alpha(x)^2 \Big| N(x) = m^+, N^-(x) = m\Big\} = E\Big\{\widehat{r}_\alpha^\star(x)^2 - \widetilde{r}_\alpha(x)^2 \Big| N^-(x) = m\Big\}$. By definition, we have that $\widehat{r}_\alpha^\star(x)$ and $\widetilde{r}_\alpha(x)$ are absolutely bounded, uniformly in $x \in R_X$ and $\alpha \in A$. This gives that

$$E\Big\{\widehat{r}_\alpha^\star(x)^2 - \widetilde{r}_\alpha(x)^2 \Big| N(x) = m^+, N^-(x) = m\Big\} \le C,$$

uniformly in $x \in R_X$, $\alpha \in A$ and $m^+ \ge m$. We will use this inequality and the fact that

$$P\left(N^-(x) \le \frac{m^+}{4} \Big| N(x) = m^+\right) \le C \exp(-cnh^d),$$

uniformly in $m^+ \ge \frac{1}{2} 3^d f_X(x) nh^d$. Using these facts and (10) we conclude that

$$E\Big\{\widehat{r}_\alpha^\star(x)^2 - \widetilde{r}_\alpha(x)^2 \Big| N(x) = m^+\Big\} = O(L_n n^{-3/2} h^{-5d/4}),$$

uniformly in $x \in R_X$, $\alpha \in A$ and $\frac{1}{2} 3^d f_X(x) nh^d \le m^+ \le 2\ 3^d f_X(x) nh^d$.

Because $L_n^* n^{-3/2} h^{-5d/4} = o(n^{-1} h^{-d/2})$ by Assumption (B3), and

$$P\left(\frac{1}{2} 3^d f_X(x) nh^d \le N(x) \le 2\ 3^d f_X(x) nh^d \text{ for all } x \in R_X\right) \to 1,$$

we get the statement of the lemma.

We now come to the proof of (10). Define $\widehat{r}_\alpha^-(x) = \widehat{r}_\alpha^0(x) - \Delta_\alpha^h(x)$ with

$$\Delta_\alpha^h(x) = n^{-1/2} h^{-d/4} \frac{E\Big[K\Big(\frac{x-X_i}{h}\Big) \Delta_\alpha(X_i) f_{\varepsilon_\alpha|X}(0|X_i) \Big| i \in \mathcal{N}^-(x)\Big]}{E\Big[K\Big(\frac{x-X_i}{h}\Big) f_{\varepsilon_\alpha|X}(0|X_i) \Big| i \in \mathcal{N}^-(x)\Big]}.$$

First note that

$$\widehat{r}_\alpha^-(x) \le u m_0^{-1/2} \quad \text{if and only if} \tag{11}$$
$$\sum_{i \in \mathcal{N}^-(x)} K\Big(\frac{x-X_i}{h}\Big) \Big\{ I(\varepsilon_{i,\alpha}^\Delta \le \Delta_\alpha^h(x) + u m_0^{-1/2}) - \alpha \Big\} \ge 0.$$

Let

$$\begin{aligned} g_{x,\alpha}(u) &= E\Big[K\Big(\frac{x-X_i}{h}\Big) \Big\{ I(\varepsilon_{i,\alpha}^\Delta \le \Delta_\alpha^h(x) + u m_0^{-1/2}) - \alpha \Big\} \Big| i \in \mathcal{N}^-(x)\Big] \\ &= u m_0^{-1/2} E\Big[K\Big(\frac{x-X_i}{h}\Big) f_{\varepsilon_\alpha|X}(0|X_i) \Big| i \in \mathcal{N}^-(x)\Big] \\ &\quad + \frac{1}{2} u^2 m_0^{-1} E\Big[K\Big(\frac{x-X_i}{h}\Big) f'_{\varepsilon_\alpha|X}(0|X_i) \Big| i \in \mathcal{N}^-(x)\Big] \\ &\quad + O(L_n n^{-1} h^{-3d/4}), \end{aligned}$$

uniformly in $|u| \leq C^* L_n^*$, because of Assumption (B3). Then, with

$$\eta_{i,\alpha,u,x} = K\Big(\frac{x - X_i}{h}\Big)\Big\{I(\varepsilon_{i,\alpha}^{\Delta} \leq \Delta_{\alpha}^{h}(x) + u m_0^{-1/2}) - \alpha\Big\} - g_{x,\alpha}(u),$$

we have that

$$P\Big(\widehat{r_{\alpha}^{-}}(x) \leq u m_0^{-1/2}\Big|\mathcal{N}^{-}(x), N^{-}(x) = m\Big)$$
$$= P\Big(m^{-1/2}\sum_{i\in\mathcal{N}^{-}(x)}\eta_{i,\alpha,u,x} \geq -m^{1/2}g_{x,\alpha}(u)\Big|\mathcal{N}^{-}(x), N^{-}(x) = m\Big).$$

We now argue that an Edgeworth expansion holds for the density of $m^{-1/2}\sum_{i\in\mathcal{N}^{-}(x)}\eta_{i,\alpha,u,x}$ that is of the form

$$\sigma^{-1}\sum_{r=0}^{S-3}P_r(-\phi : \{\bar{\chi}_{\beta,r}\})(\sigma^{-1}[\cdot - x - \mu_{nn}]) + O(n^{-(S-2)/2}[1 + |\sigma^{-1}[\cdot - x - \mu_{nn}]|^{S}]^{-1})$$

with standard notations, see Bhattacharya and Rao (1976), p. 53. In particular, P_r denotes a product of a standard normal density with a polynomial that has coefficients depending only on cumulants of order $\leq r + 2$. In our case such an expansion follows from Theorem 19.3 in Bhattacharya and Rao (1976). For this claim we have to verify that their conditions (19.27), (19.29) and (19.30) hold. Our setting is slightly different from theirs, since we consider triangular arrays of independent identically distributed random variables instead of a sequence of independent random variables as is the case in Theorem 19.3 in Bhattacharya and Rao (1976). But the same proof applies because in our setting we can verify uniform versions of (19.27), (19.29) and (19.30). This can be directly seen for (19.27). For checking (19.29), we consider the conditional density of $U_p = \sum_{j=1}^{p}(k(\frac{x_1 - X_{1,j}}{h}), ..., k(\frac{x_d - X_{d,j}}{h}))\Big\{I(\varepsilon_{j,\alpha}^{\Delta} \leq \Delta_{\alpha}^{h}(x) + u m_0^{-1/2}) - \alpha\Big\}$ given the value of $\varepsilon_{j,\alpha}^{\Delta}$ and given that $X_j \in \mathcal{N}^{-}(x)$ for $j = 1, ..., p$. For $p = 1$ this density can be bounded by a constant times $(u_1 \cdot ... \cdot u_d)^{-\kappa}((k(0) - u_1) \cdot ... \cdot (k(0) - u_d))^{-\kappa}$ by Assumptions (B1) and (B3). This bound holds uniformly over α, u, x and the value of $\varepsilon_{1,\alpha}^{\Delta}$. Furthermore, for $p > \kappa$ we get that the conditional density of U_p is uniformly bounded. From this we conclude that the conditional density of $\sum_{j=1}^{p}K(\frac{x - X_j}{h})\Big\{I(\varepsilon_{j,\alpha}^{\Delta} \leq \Delta_{\alpha}^{h}(x) + u m_0^{-1/2}) - \alpha\Big\} = \sum_{j=1}^{p}k(\frac{x_1 - X_{1,j}}{h}) \cdot ... \cdot k(\frac{x_d - X_{d,j}}{h})\Big\{I(\varepsilon_{j,\alpha}^{\Delta} \leq \Delta_{\alpha}^{h}(x) + u m_0^{-1/2}) - \alpha\Big\}$ is uniformly bounded and thus also the same holds for the conditional density of $\sum_{j=1}^{p} \eta_{j,\alpha,u,x}$, given the value of $\varepsilon_{j,\alpha}^{\Delta}$ and given that $X_j \in \mathcal{N}^-(x)$ for $j = 1, ..., p$. Now, this gives that also the conditional density of $\sum_{j=1}^{p} \eta_{j,\alpha,u,x}$ is bounded, given that $X_j \in \mathcal{N}^-(x)$ for $j = 1, ..., p$, uniformly over α, u and x. Thus, the square of this conditional density is integrable and by the Fourier Inversion Theorem (see Theorem 4.1 (vi) in Bhattacharya and Rao (1976)) the same holds for the squared modulus of its Fourier transform. Thus the modulus of the Fourier transform of the conditional density of $\sum_{j=1}^{2p} \eta_{j,\alpha,u,x}$, given that $X_j \in \mathcal{N}^-(x)$ for $j = 1, ..., 2p$, is integrable. This shows (19.29). For the proof of (19.30) one applies the Riemann-Lebesgue Lemma (see Theorem 4.1 in Bhattacharya and Rao (1976)). Consider for simplicity the case where $d = 1$. From the Riemann-Lebesgue Lemma we get that $\sup_{|t|\geq b} \left| \int_{v_0}^{v_1} \exp(itK(v))\, dv \right| < v_1 - v_0$ for $0 < v_0 < v_1$ with $b > 0$ large enough. Furthermore, one can approximate $E[\exp(it\eta_{j,\alpha,u,x})|X_j \in \mathcal{N}^-(x)]$ by

$$\frac{1}{3} \int_{-2+\rho(x)}^{1+\rho(x)} \int_{e\in\mathbb{R}} \exp[itK(v)\{\mathrm{I}(e \leq \Delta_\alpha^h(x) + um_0^{-1/2} - n^{-1/2}h^{-d/4}\Delta_\alpha(x)) - \alpha\}] \\ \times \exp[-itg_{x,\alpha}(u)]\; f_{\varepsilon_\alpha|X}(e|x)\; dv\; de$$

for some function $0 \leq \rho(x) \leq 1$. Furthermore, using our assumptions this can be approximated by

$$\frac{\alpha}{3} \int_{-2+\rho(x)}^{1+\rho(x)} \exp[it(1-\alpha)K(v)]dv + \frac{1-\alpha}{3} \int_{-2+\rho(x)}^{1+\rho(x)} \exp[-it\alpha K(v)]dv.$$

These approximations hold uniformly in $\alpha \in A$, $|u| \leq L_n^* m_0^{-1/2}$ and $x \in R_X$. By using these facts we get (19.30).

By applying Theorem 19.3 in Bhattacharya and Rao (1976) with $s \geq 4$ we get that

$$P\Big(\widehat{r_\alpha^-}(x) \leq um_0^{-1/2} \Big| \mathcal{N}^-(x), N^-(x) = m\Big) \tag{12}$$
$$= 1 - \Phi\Big(\mu_\alpha(u)\Big) + m^{-1/2}\rho_\alpha(u)\Big(1 - \mu_\alpha(u)^2\Big)\phi\Big(\mu_\alpha(u)\Big) + O\Big(m_0^{-1}(1 + \mu_\alpha(u)^2)^{-s}\Big),$$

uniformly in u, α and x for $C_1^* m_0 \leq m \leq C_2^* m_0$ and constants $C_1^* < C_2^*$. Here we have used the fact that terms for $r = 2, ..., s-3$ in the expansion of Theorem 19.3 in

Bhattacharya and Rao (1976) can be bounded by $O\Big(m_0^{-1}(1+\mu_\alpha(u)^2)^{-s}\Big)$. We used the following notation

$$\mu_\alpha(u) = -\frac{m^{1/2} g_{x,\alpha}(u)}{\sigma_\alpha(u)} \quad \text{and} \quad \rho_\alpha(u) = \frac{m^{-1}\sum_{i=1}^n E(\eta_{i,\alpha,u,x}^3 | N^-(x) = m)}{\sigma_\alpha^3(u)},$$

with $\sigma_\alpha^2(u) = m^{-1}\sum_{i=1}^n E(\eta_{i,\alpha,u,x}^2 | N^-(x) = m)$. It is easy to show that, uniformly in $|u| \le C^* L_n^*$,

$$\begin{aligned}
&\sigma_\alpha^2(u) \\
&= \alpha(1-\alpha)E\Big[K^2\Big(\frac{x-X_i}{h}\Big)\Big\{1 + (um_0^{-1/2} - n^{-1/2}h^{-d/4}\Delta_\alpha(X_i))f_{\varepsilon_\alpha|X}(0|X_i)\Big\}\Big| i \in \mathcal{N}^-(x)\Big] \\
&\qquad + O(L_n m_0^{-1}),
\end{aligned}$$

and that

$$\begin{aligned}
&m^{-1}\sum_{i=1}^n E(\eta_{i,\alpha,u,x}^3 | N^-(x) = m) \\
&= E\Big[K^3\Big(\frac{x-X_i}{h}\Big)\Big| i \in \mathcal{N}^-(x)\Big]\alpha(1 - 3\alpha + 2\alpha^2) + O(L_n m_0^{-1/2}).
\end{aligned}$$

Note that $\mu_\alpha(-u) = \mu_\alpha(u) + O(L_n n^{-1/2}h^{-d/4})$, where we used that $L_n^* n^{-1}h^{-d} = O(L_n n^{-1/2}h^{-d/4})$ because of Assumption (B3). Note also that with $u_m = um^{1/2}m_0^{-1/2}$, uniformly in $|u| \le C^* L_n^*$,

$$\begin{aligned}
&1 - \Phi\Big(\mu_\alpha(u)\Big) \\
&= 1 - \Phi\Big(\frac{u_m A_n}{B_n^{1/2}}\Big) + \phi\Big(\frac{u_m A_n}{B_n^{1/2}}\Big)\frac{u_m^2 A_n}{2B_n^{3/2}}\alpha(1-\alpha) \\
&\qquad \times m^{-3/2}\sum_{i=1}^n E\Big[K^2\Big(\frac{x-X_i}{h}\Big)f_{\varepsilon_\alpha|X}(0|X_i)\Big| N^-(x) = m\Big] \\
&\quad + \phi\Big(\frac{u_m A_n}{B_n^{1/2}}\Big)\frac{u_m^2}{2B_n^{1/2}}m^{-3/2}\sum_{i=1}^n E\Big[K\Big(\frac{x-X_i}{h}\Big)f'_{\varepsilon_\alpha|X}(0|X_i)\Big| N^-(x) = m\Big] \\
&\quad + O(L_n n^{-1/2}h^{-d/4}),
\end{aligned}$$

where

$$A_n = m^{-1}\sum_{i=1}^n E\Big[K\Big(\frac{x-X_i}{h}\Big)f_{\varepsilon_\alpha|X}(0|X_i)\Big| N^-(x) = m\Big]$$

and

$$B_n = \alpha(1-\alpha)m^{-1}\sum_{i=1}^{n} E\Big[K^2\Big(\frac{x-X_i}{h}\Big)\Big|N^-(x)=m\Big].$$

Hence, uniformly in $|u| \le C^* L_n^*$,

$$1-\Phi(\mu_\alpha(u))+\Phi(-\mu_\alpha(-u)) = 2\Big[1-\Phi\Big(\frac{u_m A_n}{B_n^{1/2}}\Big)\Big] + O(L_n n^{-1/2}h^{-d/4}). \qquad (13)$$

From (13) and the above calculations it now follows that

$$\begin{aligned}
&E\Big\{\widehat{r_\alpha^-}(x)^2 I(|\widehat{r_\alpha^-}(x)| \le L_n^* m_0^{-1/2})\Big|N^-(x)=m\Big\} \\
&= 2m_0^{-1}\int_0^{L_n^*} vP\Big(\widehat{r_\alpha^-}(x) > vm_0^{-1/2}\Big|N^-(x)=m\Big)\,dv \\
&\qquad -2m_0^{-1}\int_{-L_n^*}^{0} vP\Big(\widehat{r_\alpha^-}(x) \le vm_0^{-1/2}\Big|N^-(x)=m\Big)\,dv \\
&= 2m_0^{-1}\int_0^{L_n^*} v\Big[P\Big(\widehat{r_\alpha^-}(x) > vm_0^{-1/2}\Big|N^-(x)=m\Big) + P\Big(\widehat{r_\alpha^-}(x) \le -vm_0^{-1/2}\Big|N^-(x)\Big)\Big]\,dv \\
&= 4m_0^{-1}\int_0^{L_n^*} v\Big[1-\Phi\Big(\frac{vm^{1/2}m_0^{-1/2}A_n}{B_n^{1/2}}\Big)\Big]\,dv + O(L_n n^{-3/2}h^{-5d/4}) \\
&= 4m^{-1}\int_0^{L_n^*} v\Big[1-\Phi\Big(\frac{vA_n}{B_n^{1/2}}\Big)\Big]\,dv + O(L_n n^{-3/2}h^{-5d/4}) \\
&= 2m^{-1}\Big[(L_n^*)^2\Big\{1-\Phi\Big(\frac{L_n^*A_n}{B_n^{1/2}}\Big)\Big\} + \int_0^{L_n^*} v^2\frac{A_n}{B_n^{1/2}}\phi\Big(\frac{vA_n}{B_n^{1/2}}\Big)\,dv\Big] + O(L_n n^{-3/2}h^{-5d/4}) \\
&= 2m^{-1}\frac{B_n}{A_n^2}\int_0^{L_n^*A_nB_n^{-1/2}} z^2\phi(z)\,dz + O(L_n n^{-3/2}h^{-5d/4}),
\end{aligned}$$

uniformly in $C_1^* m_0 \le m \le C_2^* m_0$ with constants $C_1^* < C_2^*$. If $L_n^* = (\log n)^\gamma$ is chosen with $\gamma > 0$ large enough we get that the right hand side of the last equation is equal to $m^{-1}\frac{B_n}{A_n^2} + O_P(L_n n^{-3/2}h^{-5d/4})$. This follows since it can be easily shown that

$$2\int_0^{L_n^*A_nB_n^{-1/2}} z^2\phi(z)\,dz - 1 = o(L_n n^{-C}) = o(m^{-2}).$$

Next, consider

$$E\Big\{(\widetilde{r}_\alpha(x)-\Delta^h_\alpha(x))^2\Big|N^-(x)=m\Big\}$$
$$=E\Big\{\Big(m^{-1}\sum_{i=1}^n K\Big(\frac{x-X_i}{h}\Big)f_{\varepsilon_\alpha|X}(0|X_i)\Big)^{-2}m^{-2}\sum_{i=1}^n K^2\Big(\frac{x-X_i}{h}\Big)\Big|N^-(x)=m\Big\}\alpha(1-\alpha)$$
$$+O(L_n n^{-2}h^{-3d/2})$$
$$=m^{-1}\frac{B_n}{A_n^2}+O(L_n n^{-3/2}h^{-5d/4}).$$

The last two expansions give that

$$E\Big\{\widehat{r_\alpha^-}(x)^2 I(|\widehat{r_\alpha^-}(x)|\le L_n m_0^{-1/2})\Big|N^-(x)=m\Big\} \tag{14}$$
$$=E\Big\{(\widetilde{r}_\alpha(x)-\Delta^h_\alpha(x))^2\Big|N^-(x)=m\Big\}+O(L_n n^{-3/2}h^{-5d/4}),$$

uniformly in $x\in R_X$, $\alpha\in A$ and $C_1^*m_0\le m\le C_2^*m_0$. Similarly one can show that

$$E\Big\{\widehat{r_\alpha^-}(x)I(|\widehat{r_\alpha^-}(x)|\le L_n^*m_0^{-1/2})\Big|N^-(x)=m\Big\}$$
$$=2m_0^{-1/2}\int_0^{L_n^*}P\Big(\widehat{r_\alpha^-}(x)>vm_0^{-1/2}\Big|N^-(x)=m\Big)\,dv$$
$$-2m_0^{-1/2}\int_{-L_n^*}^0 P\Big(\widehat{r_\alpha^-}(x)\le vm_0^{-1/2}\Big|N^-(x)=m\Big)\,dv$$
$$=2m_0^{-1/2}\int_0^{L_n^*}\Big[P\Big(\widehat{r_\alpha^-}(x)>vm_0^{-1/2}\Big|N^-(x)=m\Big)-P\Big(\widehat{r_\alpha^-}(x)\le -vm_0^{-1/2}\Big|N^-(x)\Big)\Big]\,dv$$
$$=O(L_n n^{-1}h^{-d})$$

and

$$E\Big\{\widetilde{r}_\alpha(x)-\Delta^h_\alpha(x)\Big|N^-(x)=m\Big\}=O(L_n n^{-1}h^{-d/2}).$$

The last two expansions give that

$$\Delta^h_\alpha(x)\Big[E\Big\{\widehat{r_\alpha^-}(x)I(|\widehat{r_\alpha^-}(x)|\le L_n m_0^{-1/2})\Big|N^-(x)=m\Big\}-E\Big\{\widetilde{r}_\alpha(x)-\Delta^h_\alpha(x)\Big|N^-(x)=m\Big\}\Big]$$
$$=O(L_n n^{-3/2}h^{-5d/4}), \tag{15}$$

uniformly in $x\in R_X$, $\alpha\in A$ and $C_1^*m_0\le m\le C_2^*m_0$.

From (14)–(15) we get that (10) holds. This concludes the proof of the lemma. □

Lemma 6. *Suppose the assumptions of Theorem 1 are satisfied. Then,*

$$T_{n3} = O_P(L_n n^{-5/4} h^{-3d/4}) = o_P((nh^{d/2})^{-1}).$$

Proof of Lemma 6. For simplicity of exposition of the argument, let us assume that X_i is one-dimensional. For arbitrary j and for $k = 1, 2, 3$, define

$$U_{jk} = \int_A \int_{I_{jk}} \Big[\widehat{r}^{\star}_{\alpha}(x)^2 - \widetilde{r}_{\alpha}(x)^2 - E\Big\{\widehat{r}^{\star}_{\alpha}(x)^2 - \widetilde{r}_{\alpha}(x)^2 \Big| N(x)\Big\}\Big] w(x,\alpha)\, dx\, d\alpha.$$

Then we can write $T_{n3} = T_{n31} + T_{n32} + T_{n33}$ with $T_{n3k} = \sum_j U_{jk}$ $(k = 1, 2, 3)$. The terms T_{n31}, T_{n32} and T_{n33} are sums of $O(h^{-1})$ conditionally independent summands. The summands are uniformly bounded by a term of order $O_P(L_n n^{-5/4} h^{-1/4})$. This follows from Lemma 5, from the fact that $\sup_{\alpha\in A}\sup_x |\widetilde{r}_{\alpha}(x)| = O_P(L_n(nh)^{-1/2})$, see also (8), and from the Bahadur representation for $\widehat{r}^{\star}_{\alpha}(x)$, given in Lemma 3. It now follows that $T_{n3k} = O_P(L_n n^{-5/4} h^{-3/4})$, which implies the statement of the lemma for $d = 1$. For $d > 1$ one can use the same approach. □

Lemma 7. *Suppose the assumptions of Theorem 1 are satisfied. Then,*

$$T_{n4} = o_P((nh^{d/2})^{-1}).$$

Proof of Lemma 7. This is obvious, since $T_{n4} = O_P(L_n (nh^d)^{-3/4} n^{-\frac{1}{2}+c}) = o_P((nh^{d/2})^{-1})$, thanks to Assumption (B5) and Lemma 3. □

Lemma 8. *Suppose the assumptions of Theorem 1 are satisfied. Then,*

$$T_{n5} = o_P((nh^{d/2})^{-1}).$$

Proof of Lemma 8. Write

$$\begin{aligned}
T_{n5} &= 2\int_A \int_{R_X} \frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)\{I(\varepsilon^{\Delta}_{i,\alpha} \le 0) - \alpha\}}{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) f_{\varepsilon_\alpha|X}(0|X_i)} (\widehat{\theta}(\alpha) - \theta_0(\alpha))^{\top} \gamma_\alpha(x) w(x,\alpha)\, dx\, d\alpha \\
&= \frac{2}{n}\int_A \int_{R_X} \frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)\{I(\varepsilon^{\Delta}_{i,\alpha} \le 0) - \alpha\}}{g_{h,\alpha}(x)} \\
&\qquad \times (\widehat{\theta}(\alpha) - \theta_0(\alpha))^{\top} \gamma_\alpha(x) w(x,\alpha)\, dx\, d\alpha + o_P((nh^{d/2})^{-1}) \\
&= 2\int_A (\widehat{\theta}(\alpha) - \theta_0(\alpha))^{\top} \frac{1}{n}\sum_{i=1}^n \rho_{h,\alpha}(X_i)\{I(\varepsilon^{\Delta}_{i,\alpha} \le 0) - \alpha\}\, d\alpha + o_P((nh^{d/2})^{-1}), \qquad (16)
\end{aligned}$$

with $g_{h,\alpha}(x) = E\big[K\big(\frac{x-X}{h}\big) f_{\varepsilon_\alpha|X}(0|X)\big]$ and

$$\rho_{h,\alpha}(v) = \int_{R_X} K\Big(\frac{x-v}{h}\Big) \frac{\gamma_\alpha(x) w(x,\alpha)}{g_{h,\alpha}(x)}\, dx.$$

Using the notations $Q_{h,\alpha}(X_i) = \frac{\rho_{h,\alpha}(X_i)}{\sum_{j=1}^n \rho_{h,\alpha}(X_j)}$, $\widehat{F}_{\varepsilon_\alpha^\Delta}(y) = \sum_{i=1}^n Q_{h,\alpha}(X_i) I(\varepsilon_{i,\alpha}^\Delta \le y)$ and $F_{\varepsilon_\alpha^\Delta}(y) = P(\varepsilon_\alpha^\Delta \le y)$, we have that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \rho_{h,\alpha}(X_i)\{I(\varepsilon_{i,\alpha}^\Delta \le 0) - \alpha\} \\
&= \Big[\widehat{F}_{\varepsilon_\alpha^\Delta}(0) - \alpha\Big]\Big(\frac{1}{n}\sum_{i=1}^n \rho_{h,\alpha}(X_i)\Big) \\
&= \Big[\widehat{F}_{\varepsilon_\alpha^\Delta}(0) - F_{\varepsilon_\alpha^\Delta}(0)\Big]\Big(\frac{1}{n}\sum_{i=1}^n \rho_{h,\alpha}(X_i)\Big) + \Big[F_{\varepsilon_\alpha^\Delta}(0) - \alpha\Big]\Big(\frac{1}{n}\sum_{i=1}^n \rho_{h,\alpha}(X_i)\Big) \\
&= O_P(n^{-1/2}) + O_P(n^{-1/2}h^{-d/4}),
\end{aligned}$$

uniformly in $\alpha \in A$, and hence (16) is $O_P(n^{-1+c}h^{-d/4}) = o_P((nh^{d/2})^{-1})$ for c small enough.

□

Lemma 9. *Suppose the assumptions of Theorem 1 are satisfied. Then,*

$$T_{n6} = o_P((nh^{d/2})^{-1}).$$

Proof of Lemma 9. The proof is obvious, since by Assumption (B5), we have that $T_{n6} = O_P(n^{-1+2c}) = o_P((nh^{d/2})^{-1})$ for c small enough. □

Lemma 10. *Suppose the assumptions of Theorem 1 are satisfied. Then,*

$$nh^{d/2} T_{n7} - b_{h,A} \xrightarrow{d} N(D_A, V_A).$$

Proof of Lemma 10. The proof is very similar to the proof of e.g. Proposition 1 in Härdle and Mammen (1993). Write

$$\begin{aligned}
T_{n7} \;=\; & n^{-2}\sum_{i,j}\int_A\int_{R_X} K\Big(\frac{x-X_i}{h}\Big)K\Big(\frac{x-X_j}{h}\Big)\{I(\varepsilon_{i,\alpha}^\Delta \le 0) - \alpha\}\{I(\varepsilon_{j,\alpha}^\Delta \le 0) - \alpha\} \\
& \times \widehat{g}_\alpha(x)^{-2} w(x,\alpha)\, dx\, d\alpha,
\end{aligned}$$

where $\widehat{g}_\alpha(x) = n^{-1}\sum_{i=1}^n K\Big(\frac{x-X_i}{h}\Big) f_{\varepsilon_\alpha|X}(0|X_i)$. By writing $I(\varepsilon_{i,\alpha}^\Delta \le 0) - \alpha = \big[I(\varepsilon_{i,\alpha}^\Delta \le 0) - I(\varepsilon_{i,\alpha} \le 0)\big] + \big[I(\varepsilon_{i,\alpha} \le 0) - \alpha\big]$, we can decompose T_{n7} into $T_{n7} = T_{n71} + T_{n72} + 2T_{n73}$.

As in Härdle and Mammen (1993), T_{n73} is negligible. Straightforward calculations show that $T_{n71} = (nh^{d/2})^{-1}(D_A + o_P(1))$. Next, write $T_{n72} = T_{n72a} + T_{n72b}$ with

$$T_{n72a} = \frac{1}{n^2}\sum_{i=1}^{n} U_{nii},$$

$$T_{n72b} = \frac{1}{n^2}\sum_{i\neq j} U_{nij},$$

where

$$U_{nij} = \int_A \int_{R_X} K\Big(\frac{x-X_i}{h}\Big)K\Big(\frac{x-X_j}{h}\Big)\{I(\varepsilon_{i,\alpha} \leq 0)-\alpha\}\{I(\varepsilon_{j,\alpha} \leq 0)-\alpha\} \times \widehat{g}_\alpha(x)^{-2}w(x,\alpha)\,dx\,d\alpha.$$

By calculating its mean and variance it can be checked that $nh^{d/2}T_{n72a} = b_{h,A} + o_P(1)$. Thus for the lemma it remains to check that $nh^{d/2}T_{n72b} \stackrel{d}{\to} N(0, V_A)$. For the proof of this claim one can proceed as in Härdle and Mammen (1993) and apply the central limit theorem for U-statistics of de Jong (1987). For this purpose one has to verify that $n^2h^d\mathrm{Var}(T_{n72b}) \to V_A$, $\max_{1\leq i\leq n}\sum_{j=1}^n \mathrm{Var}(U_{nij})/\mathrm{Var}(T_{n72b}) \to 0$ and $E[T_{n72b}^4]/(\mathrm{Var}(T_{n72b}))^2 \to 3$. This can be done by straightforward but tedious calculations. □

Proof of Theorem 1. The theorem follows immediately from Lemmas 4–10. Lemmas 4–9 imply the negligibility of the terms T_{n1}, .., T_{n6}. Lemma 10 shows the asymptotic normality of $nh^{d/2}T_{n7}$. □

4 Proof of Theorem 2

The theorem can be shown by verification of the conditions of the central limit theorem for U-statistics of de Jong (1987), in the same way as was done in the proof of Lemma 10. The crucial point in the proof is to note that $I(U_i \leq \alpha)$ has the same distribution as $I(\varepsilon_{i,\alpha} \leq 0)$, and hence the calculations in the proof of Lemma 10 go through in this proof.

5 Proof of Theorem 3

For the proof of Theorem 3 one can proceed similarly as in the proof of Theorem 1. Using Theorem 2 in Guerre and Sabbah (2012) we get a Bahadur expansion of $\widehat{r}_{\alpha}^{\dagger}(x)$, of which the remainder term is of the order $O_P(L_n(nh^d)^{-3/4})$. Because now $L_n(nh^d)^{-3/4} \times (nh^d)^{-1/2}$ is of lower order than $O(n^{-1}h^{-d/2})$ one can directly replace $\widehat{r}_{\alpha}^{\dagger}(x)$ in the definition of $\widehat{T}_A^{\dagger}$ by its Bahadur expansion and proceed as in the proof of Lemma 10.

Acknowledgments

The first author was financially supported by the Deutsche Forschungsgemeinschaft (German Science Foundation) in the framework of the Swiss-German Research Group 'Statistical Regularisation and Qualitative Constraints'. The research of the second author was supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement No. 203650, by IAP research network grant nr. P7/06 of the Belgian government (Belgian Science Policy), and by the contract "Projet d'Actions de Recherche Concertées" (ARC) 11/16-039 of the "Communauté française de Belgique" (granted by the "Académie universitaire Louvain").

References

Ait-Sahalia, Y., Fan, J. and Peng, H. (2009). Nonparametric transition-based tests for diffusions. *Journal of the American Statistical Association* **104** 1102-1116.

Bhattacharya, R. and Rao, R. (1976). *Normal Approximations and Asymptotic Expansions.* John Wiley & Sons, New York.

Chaudhuri, P. (1991). Nonparametric estimates of regression quantiles and their local Bahadur representation. *Annals of Statistics* **19** 760-777.

de Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probability*

Theory and Related Fields **75** 261-277.

Dette, H. and Sprekelsen, I. (2004). Some comments on specification tests in nonparametric absolutely regular processes. *Journal of Time Series Analysis* **25** 159-172.

El Ghouch, A. and Van Keilegom, I. (2009). Local linear quantile regression with dependent censored data. *Statistica Sinica* **19** 1621-1640.

Fan, J., Zhang, C. and Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Annals of Statistics* **29** 153-193.

Gao, J. and Hong, Y. (2008). Central limit theorems for generalized U-statistics with applications in nonparametric specification. *Journal of Nonparametric Statistics* **20** 61-76.

González-Manteiga, W. and Cao-Abad, R. (1993). Testing the hypothesis of a general linear model using nonparametric regression estimation. *Test* **2** 161-188.

Guerre, E. and Lavergne, P. (2002). Optimal minimax rates for nonparametric specification testing in regression models. *Econometric Theory* **18** 1139-1171.

Guerre, E. and Sabbah, C. (2012). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. *Econometric Theory* **28** 87-129.

Haag, B. (2008). Non-parametric regression tests using dimension reduction techniques. *Scandinavian Journal of Statistics* **35** 719-738.

Härdle, W. and Mammen, E. (1993). Testing parametric versus nonparametric regression. *Annals of Statistics* **21** 1926-1947.

He, X. and Ng, P. (1999). Quantile splines with several covariates. *Journal of Statistical Planning and Inference* **75** 343-352.

He, X., Ng, P. and Portnoy, S. (1998). Bivariate quantile smoothing splines. *Journal of the Royal Statistical Society - Series B* **60** 537-550.

He, X. and Zhu, L.-X. (2003). A lack of fit test for quantile regression. *Journal of the American Statistical Association* **98** 1013-1022.

Hjellvik, V., Yao, Q. and Tjøstheim, D. (1998). Linearity testing using local polynomial approximation. *Journal of Statistical Planning and Inference* **68** 295-321.

Hoderlein, S. and Mammen, E. (2009). Identification and estimation of local average derivatives in non-separable models without monotonicity. *Econometrics Journal* **12** 1-25.

Hong, S.Y. (2003). Bahadur representation and its applications for local polynomial estimates in non-parametric M-regression. *Journal of Nonparametric Statistics* **15** 237-251.

Ingster, Y.I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives I, II, III. *Math. Methods of Statistics* **2** 85-114, 171-189, 249-268.

Koenker, R. (2005). *Quantile Regression.* Cambridge University Press.

Koenker, R. and Machado, J.A.F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association* **94** 1296-1310.

Koenker, R. and Xiao, Z. J. (2002). Inference on the quantile regression process. *Econometrica* **70** 1583-1612.

Kong, E., Linton, O. and Xia, Y. (2010). Uniform Bahadur representation for local polynomial estimates of M-regression and its application to the additive model. *Econometric Theory* **26** 1529- 1564.

Kreiss, J.P., Neumann, M.H. and Yao, Q. (2008). Bootstrap tests for simple structures in nonparametric time series regression. *Statistics and its Interface* **1** 367-380.

Lee, K.L. and Lee, E.R. (2008). Kernel methods for estimating derivatives of conditional quantiles. *Journal of the Korean Statistical Society* **37** 365-373.

Leucht, A. (2012). Degenerate U- and V-statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli* **18** 552-585.

Li, Q. and Racine, J.S. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business & Economic Statistics* **26** 423-434.

Truong, Y.K. (1989). Asymptotic properties of kernel estimators based on local medians. *Annals of Statistics* **17** 606-617.

Zheng, J.X. (1996). A consistent test of a functional form via nonparametric estimation techniques. *Journal of Econometrics* **75** 263-289.

Zheng, J.X. (1998). A consistent nonparametric test of parametric models under conditional quantile regressions. *Econometric Theory* **14** 223-238.