Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 05–09

NONPARAMETRIC DENSITY AND REGRESSION ESTIMATION FOR FUNCTIONAL DATA

PAULO EDUARDO OLIVEIRA

ABSTRACT: We consider kernel estimation of density and regression based on functional data. We prove the strong convergence and the asymptotic normality of the centered estimators. We include results both for independent and mixing data, as the mathematical treatment and conditions for convergence are different.

KEYWORDS: functional data, density estimation, regression estimation, asymptotic normality.
AMS SUBJECT CLASSIFICATION (2000): 62G05, 62G08, 62M09.

1. Introduction

The problem of approximating density or regression functions is in data analysis, and there exists an extensive literature both for independent and dependent data. These problems are usually addressed in $\mathbb{R}^d$, where the Lebesgue measure plays an essential role, particularly for the density estimation, but it also appears indirectly in some approaches of regression estimation. In infinite dimensional spaces there is no Lebesgue measure nor any analogue that may be suitable to replace it. As a matter of fact, the invariance under translations of the Lebesgue measure plays an important role in the treatment of the kernel estimators. An abstracts measure analogue with this property would be a Haar measure, but this does not generally exists in an infinite dimensional space, even if it is a Hilbert space. So, even mathematically, there are some intrinsic features to be addressed when approaching these estimation problems in infinite dimensional spaces. Moreover, in recent years, there has been increasing interest in estimation based on functional data (see Ramsay, Silverman [17] for some case studies), so this framework has received more attention from the statisticians: for example, Ferraty, Vieu [6, 7] considered regression estimation and time series prediction for dependent functional data, Dabo-Niang [5] studied density estimation in Banach space with an application to the estimation of the

Received May 20, 2005.
Author partially supported by Centro de Matemática da Universidade de Coimbra, Fundação para a Ciência e Tecnologia (FCT) and POCTI

density of a diffusion process with respect to the Wiener measure, here for independent functional data, or Masry [16] who proved the mean square convergence and the asymptotic normality for the regression for strong mixing functional data. The above mentioned articles always considered kernel estimators. There exist older approaches to these problems using histograms is abstract metric spaces: Geffroy [10] for the regression with independent data, and later Jacob, Oliveira [13, 15] using an approach through point processes with independent data for the first reference and associated functional data for the second, and BenSaïd, Oliveira [2] using the same framework but for strong mixing data. In [2, 13, 15] the approach used dealt simultaneously with density estimation and regression estimation.

The present article uses the approach suggested by the point process framework of [2, 13, 15] to prove the almost complete convergence and the asymptotic normality of the kernel estimator for the density and also for the regression. This means proving the asymptotic unbiasedness of the estimators and prove the convergence to zero of the centered estimator. For the regression this is complemented with a suitable decomposition of the quotient defining the estimator, that reduces the analysis to the treatment of two terms, one of which is a density estimator and the other is, in fact, quite similar. This decomposition has been used in Bensaïd, Fabre [1], Bensaïd, Oliveira [2], Ferrieux [8] and Jacob, Oliveira [13, 15]. We assume some mild conditions on the kernel function, which does not have to be a density but will be supposed nonnegative, some regularity on conditional first and second moments, and a suitable representation of the distribution of the functional variables. When dealing with dependent samples, we assume the strong mixing coefficients decrease polynomially with convenient decrease rate. The almost complete convergence is proved under assumptions that are similar to those known for the finite dimensional framework, both for independent or mixing samples. The asymptotic normality assumes a Lindeberg hypothesis, for independent samples, and a convenient decrease rate on the mixing coefficients, that depends only on the behaviour of a volume parameter associated to the representation of the distribution of the functional variables.

2. Definitions and assumptions

On the sequel, let (X_i, Y_i), $i \geq 1$, be a random process of equally distributed random elements, where the variables Y_i, $i \geq 1$, are nonnegative real valued, and the variables X_i, $i \geq 1$, take values in some normed space $\mathbf{S}$. Let as define, for each $x \in \mathbf{S}$, the estimators

$$\widehat{f}_n(x) = \frac{1}{n\phi(h)} \sum_{i=1}^{n} K\left(\frac{\|x - X_i\|}{h}\right),$$

$$\widehat{g}_n(x) = \frac{1}{n\phi(h)} \sum_{i=1}^{n} Y_i K\left(\frac{\|x - X_i\|}{h}\right),$$

where K is a real valued function, h is bandwidth parameter depending on n (we should have written h_n, but we choose to drop the subscript for simplicity of the notation) and ϕ is a function to be described later. The analogy with the $\mathbb{R}^d$ framework is obvious, and makes clear the role this function ϕ plays. Further, let

$$\widehat{r}_n(x) = \frac{\widehat{g}_n(x)}{\widehat{f}_n(x)}.$$

This estimator, that does no longer depend on ϕ, is the Nadaraya-Watson estimator proposed by Ferraty, Vieu [7], and also studied by Masry [16].

The following lemma is used to separate the numerator and the denominator of $\widehat{r}_n(x)$.

Lemma 2.1 (Jacob, Niéré [12]). *Let Z_1 and Z_2 be nonnegative integrable random variables. Then, for $\varepsilon > 0$ small enough,*

$$\left\{\left|\frac{Z_1}{Z_2} - \frac{\mathbb{E}Z_1}{\mathbb{E}Z_2}\right| > \varepsilon\right\} \subset \left\{\left|\frac{Z_1}{\mathbb{E}Z_1} - 1\right| > \frac{\varepsilon}{4}\frac{\mathbb{E}Z_2}{\mathbb{E}Z_1}\right\} \cup \left\{\left|\frac{Z_2}{\mathbb{E}Z_2} - 1\right| > \frac{\varepsilon}{4}\frac{\mathbb{E}Z_2}{\mathbb{E}Z_1}\right\}.$$

Using the lemma it follows that, for $\varepsilon > 0$ small enough,

$$\begin{aligned} &\left\{\left|\frac{\widehat{g}_n(x)}{\widehat{f}_n(x)} - \frac{\mathbb{E}\widehat{g}_n(x)}{\mathbb{E}\widehat{f}_n(x)}\right| > \varepsilon\right\} \\ &\qquad \subset \left\{|\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)| > \frac{\varepsilon}{4}\mathbb{E}\widehat{f}_n(x)\right\} \\ &\qquad\qquad \cup \left\{\left|\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right| > \frac{\varepsilon}{4}\frac{(\mathbb{E}\widehat{f}_n(x))^2}{\mathbb{E}\widehat{g}_n(x)}\right\}. \end{aligned} \tag{1}$$

This means that, as the almost complete convergence is regarded, it is enough to establish the almost complete convergence to zero of $\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)$ and $\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)$.

We now introduce the main assumptions to be used throughout the article.

(K1): The function K is bounded nonnegative with support $[0,1]$; we denote $M = \sup_{u\in[0,1]} K(u)$.

(D1): There exist functions $\phi(h)$, differentiable and satisfying $\lim_{h\to 0}\phi(h) = 0$, and $f(x)$ such that $\mathrm{P}(\|x - X\| \le h) = f(x)\phi(h)$.

Assumption **(D1)**, without the differentiability of ϕ, has been used in Gasser, Hall, Presnell [9], Masry [16] and, in a stronger form, in Ferraty, Vieu [7]. The function $f(x)$ is interpreted as a probability density, while ϕ may be interpreted as a volume parameter. This representation holds when $\mathbf{S} = I\!\!R^d$, where $\phi(h) \sim h^d$.

In order to prove the asymptotic unbiasedness of the estimators we need to assume the following, linking the kernel and the volume parameter.

(K2):

$$\frac{h}{\phi(h)} \int_{[0,1]} K(u)\phi'(uh)\, du \longrightarrow 1,$$

$$\frac{h}{\phi(h)} \int_{[0,1]} K^2(u)\phi'(uh)\, du \longrightarrow c_2.$$

The fact the we require the convergence to 1 in the first of these two conditions, is just a matter of a convenient normalization of K. Note that this assumption is independent of the dimension of the space. This means that our results, remaing true for finite dimensional spaces, permit different normalizations of the kernel estimator other than the usual ones. It is easy to check that these assumptions on K can deal with a polynomially decreasing ϕ, as is the case for $I\!\!R^d$ with the euclidean norm, but can also resist to exponentially decreasing ϕ, as it might happen in infinite dimension.

Regarding the conditional moments of the variable Y we will assume the following.

(R1): The functions $r(x) = I\!\!E(Y|X=x)$ and $g_2(x) = I\!\!E(Y^2|X=x)$ are continuous.

(M1): There exists a constant $M_1 > 0$ such that, for every $l \ge 2$,

$$I\!\!E(Y^l|X=x) \le M_1^l l! g_2(x).$$

Assumption **(M1)** is a conditional Cramer hypothesis. It has been used in the context of regression in Geffroy [11], for example.

We start by stating and proving some asymptotic results for the first and second moments of $\widehat{f}_n(x)$. This first lemma will play the role of the classical Bochner Lemma in finite dimension.

Lemma 2.2. *Suppose* **(K1)**, **(K2)** *and* **(D1)** *hold. Then*

$$\mathbb{E}\widehat{f}_n(x) = \frac{1}{\phi(h)}\mathbb{E}K\left(\frac{\|x-X\|}{h}\right) \longrightarrow f(x),$$

$$\frac{1}{\phi(h)}\mathbb{E}K^2\left(\frac{\|x-X\|}{h}\right) \longrightarrow c_2 f(x).$$

If we suppose further that

$$n\phi(h) \longrightarrow +\infty, \tag{2}$$

then

$$\operatorname{Var}\left(\widehat{f}_n(x)\right) \longrightarrow 0.$$

Proof: Write the mathematical expectations as integrals over $[0,1]$. Then

$$\frac{1}{\phi(h)}\mathbb{E}K\left(\frac{\|x-X\|}{h}\right) = \frac{h}{\phi(h)}\int_{[0,1]} K(z)f(x)\phi'(zh)\,dz,$$

from which the result is immediate. The other convergence follows analogously and the convergence of the variance follows now readily. ■

The next lemma states similar results concerning $\widehat{g}_n(x)$.

Lemma 2.3. *Suppose* **(K1)**, **(K2)**, **(D1)** *and* **(R1)** *hold. Then*

$$\mathbb{E}\widehat{g}_n(x) = \frac{1}{\phi(h)}\mathbb{E}\left[YK\left(\frac{\|x-X\|}{h}\right)\right] \longrightarrow r(x)f(x),$$

$$\frac{1}{\phi(h)}\mathbb{E}\left[Y^2K^2\left(\frac{\|x-X\|}{h}\right)\right] \longrightarrow c_2 g_2(x)f(x).$$

If we suppose further that (2) is satisfied, then

$$\operatorname{Var}\left(\widehat{g}_n(x)\right) \longrightarrow 0.$$

Proof: The proof follows the same arguments as the proof of the previous lemma, by writing

$$\begin{aligned}\mathbb{E}\widehat{g}_n(x) &= \frac{1}{\phi(h)}\mathbb{E}\left[r(X)K\left(\frac{\|x-X\|}{h}\right)\right] \\ &= \frac{1}{\phi(h)}\int r(x)K\left(\frac{\|x-u\|}{h}\right)\,\mathrm{P}_X(du) \\ &\quad +\frac{1}{\phi(h)}\int (r(x)-r(u))K\left(\frac{\|x-u\|}{h}\right)\,\mathrm{P}_X(du),\end{aligned}$$

and remarking that

$$\begin{aligned}&\frac{1}{\phi(h)}\int (r(x)-r(u))K\left(\frac{\|x-u\|}{h}\right)\,\mathrm{P}_X(du) \\ &\leq \sup_{\|x-u\|\leq h}|r(x)-r(u)|\,\frac{1}{\phi(h)}\int K\left(\frac{\|x-u\|}{h}\right)\,\mathrm{P}_X(du)\longrightarrow 0,\end{aligned}$$

as r is continuous. ∎

These two lemmas imply the asymptotic unbiasedness of $\widehat{r}_n(x)=\frac{\widehat{g}_n(x)}{\widehat{f}_n(x)}$.

3. Asymptotics for independent samples

Suppose throughout this section, that the random elements (X_i,Y_i), $i\geq 1$, are independent. We prove the almost complete convergence, and the asymptotic normality of the estimators.

In order to prove the almost complete convergence we will apply a convenient exponential inequality. For easier reference we quote here its general form. For details we refer the reader to Lemma 1 in Jacob, Oliveira [14]

Lemma 3.1. *Let Z be a nonnegative valued random variable with finite Laplace transform on $[-\delta,\delta]$. Suppose Z_1, $Z_2,\ldots$ are independent copies of Z. Then, for every $u>0$,*

$$\mathrm{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}Z_i-\mathbb{E}Z\right|>u\right)\leq 2\exp\left(-\frac{ntu}{2}\right), \qquad (3)$$

where $t=\min\left(\delta,\frac{u}{2c}\right)$ and $c=\sum_{l=2}^{\infty}\frac{\delta^{l-2}}{l!}\mathbb{E}Z^l$.

Note that assumption **(M1)** implies the existence of Laplace transforms of each term intervening in the definition of $\widehat{g}_n(x)$, as K is supposed bounded.

3.1. Almost complete convergence. We will prove the almost complete convergence to zero of $\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)$, which corresponds to the estimation of the density, and $\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)$, from which follows the convergence of $\widehat{r}_n(x)$, using the inclusion (1).

Theorem 3.2. *Suppose the functional density $f(x) > 0$, that* **(K1)**, **(K2)**, **(D1)**, **(R1)** *and* **(M1)** *hold. If*

$$\frac{n\phi(h)}{\log n} \longrightarrow +\infty, \tag{4}$$

then $\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)$ and $\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)$ both converge almost completely to zero. The same holds for $\widehat{r}_n(x) - \mathbb{E}\widehat{r}_n(x)(x)$.

Proof: Note that the Laplace transforms of $K\left(\frac{\|x-X\|}{h}\right)$ and $YK\left(\frac{\|x-X\|}{h}\right)$ both exist on any interval of the form $[-\delta, \delta]$, as follows from **(K1)** and **(M1)**. Now, writing

$$\begin{aligned} &\mathrm{P}\left(\left|\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right| > \varepsilon\right) \\ &\quad = \mathrm{P}\left(\frac{1}{n}\left|\sum_{i=1}^n K\left(\frac{\|x-X_i\|}{h}\right) - \mathbb{E}K\left(\frac{\|x-X_i\|}{h}\right)\right| > \varepsilon\phi(h)\right), \end{aligned}$$

we may apply Lemma 3.1, where, for n large enough, we will have $t = \frac{\varepsilon\phi(h)}{2c}$, where $c = \sum_{l=2}^{\infty} \frac{\delta^{l-2}}{l!}\mathbb{E}K^l\left(\frac{\|x-X\|}{h}\right)$. As

$$\mathbb{E}K^l\left(\frac{\|x-X\|}{h}\right) = \int_{[0,h]} K^l\left(\frac{t}{h}\right) f(x)\phi'(t)\, dt \le M^l f(x)\phi(h),$$

we have $\frac{\varepsilon\phi(h)}{2c} > \frac{\varepsilon}{2c(x)f(x)}$, where $c(x) = \sum_{l=2}^{\infty} \frac{\delta^{l-2}}{l!}M^l$. These choices lead to the inequality

$$\mathrm{P}\left(\left|\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right| > \varepsilon\right) \le \exp\left(-\frac{n\varepsilon^2\phi(h)}{4c(x)f(x)}\right),$$

so the almost complete convergence follows from (4).

As what regards $\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)$ the arguments are similar, excepting the control of the argument of the exponential when applying Lemma 3.1. A direct application of Lemma 3.1 would lead to the choice of the constant

$c=\sum_{l=2}^{\infty}\frac{\delta^{l-2}}{l!}I\!\!E\left[Y^lK^l\left(\frac{\|x-X\|}{h}\right)\right]$. Now, using **(M1)** we have

$$
\begin{aligned}
&I\!\!E\left[Y^lK^l\left(\frac{\| x-X\|}{h}\right)\right]\\
&\quad=I\!\!E\left[I\!\!E(Y^l|X=x)K^l\left(\frac{\| x-X\|}{h}\right)\right]\\
&\quad\leq M_1^l l! I\!\!E\left[g_2(X)K^l\left(\frac{\| x-X\|}{h}\right)\right]\\
&\quad\leq M_1^l l!\left(g_2(x)+\sup_{\|u-x\|\leq h}|g_2(u)-g_2(x)|\right)f(x)\phi(h)\\
&\quad\leq 2M_1^l l! g_2(x)f(x)\phi(h),
\end{aligned}
$$

so the arguments of the previous case still hold, with the $c(x)$ replaced by $\frac{g_2(x)(MM_1)^2}{1-\delta MM_1}$, provided that $\delta MM_1<1$, which is always possible. ∎

The previous theorem identifies an almost sure convergence rate. In fact, it suffices to suppose that ε also depends on n and define it so that we find a convergent series. It is easily checked that the rate obtained is of order $\left(\frac{\log n}{n\phi(h)}\right)^{1/2}$.

3.2. Asymptotic normality. We will start by proving the asymptotic normality of the random vector

$$
\sqrt{n\phi(h)}\left(\widehat{f}_n(x)-I\!\!E\widehat{f}_n(x),\,\widehat{g}_n(x)-I\!\!E\widehat{g}_n(x)\right). \tag{5}
$$

The corresponding result for the estimator $\widehat{r}_n(x)$ will then follow by applying the δ-method to the function $\theta(u,v)=\frac{v}{u}$.

Theorem 3.3. *Suppose that* **(K1)**, **(K2)**, **(D1)** *and* **(R1)** *hold and that* $I\!\!EY^2<\infty$. *If, for every constant* $c>0$,

$$
\frac{1}{\phi(h)}\int_{\{Y^2>cn\phi(h)\}}Y^2\,d\mathrm{P}\longrightarrow 0, \tag{6}
$$

and

$$
n\phi^2(h)\longrightarrow+\infty, \tag{7}
$$

then, the random vector (5) converges in distribution to a centered Gaussian random vector with covariance matrix

$$\Gamma = c_2 f(x) \begin{bmatrix} 1 & r(x) \\ r(x) & g_2(x) \end{bmatrix}. \tag{8}$$

Proof: The proof consists on applying the Cramer-Wold Theorem to the random vector (5). For this purpose, let $a, b \in I\!R$, and define

$$\begin{aligned} Z_{i,n} = \frac{1}{\sqrt{n\phi(h)}} \Bigg[a \left(K\left(\frac{\|x - X_i\|}{h}\right) - I\!E K\left(\frac{\|x - X_i\|}{h}\right)\right) \\ + b\left(Y_i K\left(\frac{\|x - X_i\|}{h}\right) - I\!E\left[Y_i K\left(\frac{\|x - X_i\|}{h}\right)\right]\right)\Bigg], \end{aligned}$$

so that $\sum_{i=1}^n Z_{i,n}$ is the linear combination of the coordinates of (5) needed to apply the Cramer-Wold Theorem. We first verify that the variance of this sum is convergent. In fact,

$$\begin{aligned} &\mathrm{Var}\left(\sum_{i=1}^n Z_{i,n}\right) \\ &= \frac{1}{\phi(h)} \mathrm{Var}\left[aK\left(\frac{\|x - X\|}{h}\right) + bYK\left(\frac{\|x - X\|}{h}\right)\right] \\ &= \frac{a^2}{\phi(h)} \mathrm{Var}\left[K\left(\frac{\|x - X\|}{h}\right)\right] + \frac{b^2}{\phi(h)} \mathrm{Var}\left[YK\left(\frac{\|x - X\|}{h}\right)\right] \\ &\qquad + \frac{2ab}{\phi(h)} \mathrm{Cov}\left(K\left(\frac{\|x - X\|}{h}\right), YK\left(\frac{\|x - X\|}{h}\right)\right). \end{aligned}$$

The terms with the variances converge to $c_2 f(x)$ and $c_2 g_2(x) f(x)$, respectively, repeating the arguments used in the proofs of Lemmas 2.2 and 2.3.

As for the remaining term,

$$\begin{aligned}
&\frac{1}{\phi(h)}\operatorname{Cov}\Bigg(K\left(\frac{\|x-X\|}{h}\right), YK\left(\frac{\|x-X\|}{h}\right)\Bigg)\\
&\quad=\frac{1}{\phi(h)}\mathbb{E}\left[YK^2\left(\frac{\|x-X\|}{h}\right)\right]\\
&\qquad\qquad-\phi(h)\frac{1}{\phi(h)}\mathbb{E}\left[K\left(\frac{\|x-X\|}{h}\right)\right]\frac{1}{\phi(h)}\mathbb{E}\left[YK\left(\frac{\|x-X\|}{h}\right)\right].
\end{aligned}$$

Again, it is easy to check that the first term converges to $c_2r(x)f(x)$ while the second one converges to zero. Thus, we finally have,

$$\operatorname{Var}\left(\sum_{i=1}^{n}Z_{i,n}\right)\longrightarrow c_2f(x)\Big(a^2+b^2g_2(x)+2ab\,r(x)\Big).$$

So, to apply the Lindeberg Theorem, it is enough to verify that, for every $\varepsilon>0$,

$$\sum_{i=1}^{n}\int_{\{Z_{i,n}^2>\varepsilon^2\operatorname{Var}(\sum_{i=1}^{n}Z_{i,n})\}}Z_{i,n}^2\,d\mathrm{P}\longrightarrow 0. \tag{9}$$

Using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
Z_{i,n}^2\le\frac{4\max(a^2,b^2)}{n\phi(h)}\Bigg[&K^2\left(\frac{\|x-X_i\|}{h}\right)+\mathbb{E}^2\left(K\left(\frac{\|x-X_i\|}{h}\right)\right)\\
&+Y^2K^2\left(\frac{\|x-X_i\|}{h}\right)+\mathbb{E}^2\left(YK\left(\frac{\|x-X_i\|}{h}\right)\right)\Bigg]
\end{aligned} \tag{10}$$

Let us first simplify the integration set appearing in (9). For simplicity, put $\varepsilon'=\frac{\varepsilon^2c_2f(x)(a^2+b^2g_2(x)+2ab\,r(x))}{8\max(a^2,b^2)}$, and let $A(X_i)$ represent the expression within square brackets in (10). Then, the integration set of (9) is included in

$$\{A(X_i)>\varepsilon' n\phi(h)\}.$$

We may neglect the terms with the mathematical expectations. In fact, rewriting

$$\frac{1}{n\phi(h)}\mathbb{E}^2\left(K\left(\frac{\|x-X_i\|}{h}\right)\right)=\frac{\phi(h)}{n}\left(\frac{1}{\phi(h)}\mathbb{E}K\left(\frac{\|x-X_i\|}{h}\right)\right)^2\longrightarrow 0.$$

Analogously, we conclude that

$$\frac{1}{n\phi(h)} I\!E^2 \left(YK \left(\frac{\| x - X_i \|}{h} \right) \right) \longrightarrow 0$$

This means that the integration set of (9) is still included in

$$\left\{ K^2 \left(\frac{\| x - X_i \|}{h} \right) + Y^2 K^2 \left(\frac{\| x - X_i \|}{h} \right) > \frac{\varepsilon'}{2} n\phi(h) \right\}$$
$$\subset \left\{ K^2 \left(\frac{\| x - X_i \|}{h} \right) > \frac{\varepsilon'}{4} n\phi(h) \right\} \cup \left\{ Y^2 K^2 \left(\frac{\| x - X_i \|}{h} \right) > \frac{\varepsilon'}{4} n\phi(h) \right\}.$$

As K is bounded, and (7) implies that $n\phi(h) \longrightarrow +\infty$, the first of these two sets is, for n large enough, empty. The second one is included in

$$\left\{ Y^2 > \frac{\varepsilon'}{4M^2} n\phi(h) \right\}.$$

Now we seek for a convenient upper bound for the integrand function in (9), using again (10). For the same reason as before, the terms involving the mathematical expectations converge to zero, so we are left with

$$\frac{1}{\phi(h)} \int_{\{Y^2 > c'' n\phi(h)\}} K^2 \left(\frac{\| x - X \|}{h} \right) + Y^2 K^2 \left(\frac{\| x - X \|}{h} \right) \, d\mathrm{P},$$

where $c'' = \frac{\varepsilon'}{4M^2}$. As K is bounded by M, this integral is less or equal than

$$\frac{M^2}{\phi(h)} \mathrm{P} \left(Y^2 > c'' n\phi(h) \right) + \frac{M^2}{\phi(h)} \int_{\{Y^2 > c'' n\phi(h)\}} Y^2 \, d\mathrm{P}.$$

The second term converges to zero according to (6). Finally, using Markov's inequality,

$$\frac{1}{\phi(h)} \mathrm{P} \left(Y^2 > c'' n\phi(h) \right) \leq \frac{I\!E Y^2}{c'' n\phi^2(h)} \longrightarrow 0,$$

according to (7). ■

Applying now the δ-method to the function $\theta(u, v) = \frac{v}{u}$, we get the asymptotic normality for the regression estimator $\widehat{r}_n(x)$.

Theorem 3.4. *Suppose all the assumptions of Theorem 3.3 are satisfied. Then*

$$\sqrt{n\phi(h)} \left(\widehat{r}_n(x) - I\!E \widehat{r}_n(x) \right) \xrightarrow{d} \mathcal{N} \left(0, c_2 \left(g_2(x) - \frac{r^2(x)}{f(x)} \right) \right).$$

4. Strong mixing samples

Suppose in this section that the random elements (X_i, Y_i), $i \geq 1$, are strong mixing with coefficients $\alpha(n)$. To prove the almost complete convergence and the asymptotic normality as before, we need some further assumptions and a suitable version of the exponential inequality quoted in Lemma 3.1.

Let us introduce some assumptions needed to take care of the dependence.

(D2): There exist functions $\psi(h)$, satisfying $\lim_{h\to 0}\psi(h) = 0$, and $f_2(x)$ such that, for every distinct $i, j \geq 1$,

$$\mathrm{P}(\| x - X_i \| \leq h, \| x - X_j \| \leq h) = f_2(x)\psi(h).$$

(R2): For each $i, j \geq 1$, let $s_{ij}(u,v) = I\!E(Y_iY_j|X_i = u, X_j = v)$ and $s_{ij}^*(u,v) = I\!E(Y_j|X_i = u, X_j = v)$. Each of these set of functions is equicontinuous with respect to i, j, that is, for each $x \in \mathbf{S}$ fixed,

$$\begin{aligned}\sup_{i,j\geq 1} \sup_{\|x-u\|\leq h, \|x-v\|\leq h} |s_{ij}(u,v) - s_{ij}(x,x)| \longrightarrow 0,\\ \sup_{i,j\geq 1} \sup_{\|x-u\|\leq h, \|x-v\|\leq h} \left|s_{ij}^*(u,v) - s_{ij}^*(x,x)\right| \longrightarrow 0.\end{aligned}$$

(α1): The mixing coefficients satisfy $\alpha(n) \sim n^{-\beta}$, for some $\beta > 0$.

As before, we now quote the version of the exponential inequality we will be using. For details we refer the reader to Bensaïd, Fabre [1].

Lemma 4.1. *Let Z_i, $i \geq 1$, be a strong mixing sequence of real random variables with Laplace transforms uniformly bounded on some interval $[-\delta, \delta]$. Then, for every $n \geq 2$, $\gamma \geq 2$, $\varepsilon > 0$ and $p \leq n/2$, we have*

$$\begin{aligned}\mathrm{P}\left(\frac{1}{n}\left|\sum_{i=1}^n Z_i - I\!E Z_i\right| > \varepsilon\right)\\ \leq 6\exp\left(-\frac{nt\varepsilon}{30p}\right) + 6\frac{n}{p}\left(\frac{10M_\gamma}{\varepsilon}+1\right)^{\gamma/(2\gamma+1)}(\alpha(p))^{2\gamma/(2\gamma+1)},\end{aligned} \tag{11}$$

where $M_\gamma = \sup_{i\geq 1}\| Z_i \|_\gamma$, $t = \min\left(\frac{\delta}{2}, \frac{\varepsilon}{3c}\right)$ and $c = 4\sup_{i\geq 1}\sum_{l=2}^{\infty}\frac{\delta^{l-2}}{l!}I\!E\,|Z_i|^l$.

Finally, we quote one more general result concerning strong mixing sequences of random vectors, that will be used when proving the asymptotic normality. This a modification of the well known Bradley's coupling Lemma [4], due to Rhomari [18] (see Lemma 1.2 in Bosq [3] for an easier to find reference).

Lemma 4.2. *Let (T,W) be a $\mathbb{R}^d \times \mathbb{R}$ valued random vector such that $W \in \mathcal{L}^\theta$, for some $\theta \in [1,+\infty]$. Let $c \in \mathbb{R}$ be such that $\|W+c\|_\theta > 0$ and $u \in (0, \|W+c\|_\theta]$. Then, there exists a random variable W^* such that*

(a) *W^* has the same distribution as W and is independent of T,*

(b) $\mathrm{P}(|W^* - W| > u) \leq 11 \left(\frac{\|W+c\|_\theta}{u}\right)^{\theta/(2\theta+1)} (\alpha(\sigma(T),\sigma(W))^{2\theta/(2\theta+1)}$,

where $\alpha(\sigma(T),\sigma(W))$ is the strong mixing coefficient between the two given σ-algebras.

We may now proceed to the proof of the convergence and asymptotic normality

4.1. Almost complete convergence. The guiding line of the approach is the same as the one used for the independent case. That is, using the inclusion (1) we reduce the convergence of $\widehat{r}_n(x) - \mathbb{E}\widehat{r}_n(x)$ at the convergence of both $\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)$ and $\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)$, and use the exponential inequality of Lemma 4.1.

Theorem 4.3. *Suppose the functional density $f(x) > 0$, that* **(K1)**, **(K2)**, **(D1)**, **(R1)**, **(M1)** *and* **(α1)** *hold, with $\beta > 15/4$, for the later. If $\phi(h) \sim n^{-1+\eta}$ with $\eta > 24/(9+4\beta)$, then $\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)$ and $\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)$ converge almost completely to zero. The same holds for $\widehat{r}_n(x) - \mathbb{E}\widehat{r}_n(x)$.*

Proof: As the assumptions imply the Laplace transforms exist on every interval, we may fix δ, defined in Lemma 4.1, as it suits our needs. To apply (11) we choose $\gamma = 2$, so that $M_\gamma \leq 2M$. Also, for n large enough, $t = \frac{\varepsilon\phi(h)}{3c}$ and $c = 4\sum_{l=2}^{\infty} \frac{\delta^{l-2}}{l!} \mathbb{E}K^l\left(\frac{\|x-X\|}{h}\right) \leq 4f(x)\phi(h)\sum_{l=2}^{\infty}\frac{\delta^{l-2}M^l}{l!} =: \phi(h)c'$. Finally, in order to use the exponential inequality (11), choose $p = \sqrt{n\phi(h)}$, to find

$$\begin{aligned}
&\mathrm{P}\left(\left|\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\right| > \varepsilon\right) \\
&\quad\leq 6\exp\left(-\frac{n^{1/2}\,\varepsilon^2\phi^{1/2}(h)}{90c'}\right) + 6\frac{n^{1/2}}{\phi^{1/2}(h)}\left(\frac{10M}{\varepsilon\phi(h)}+1\right)^{2/5} n^{-2\beta/5}\phi^{-2\beta/5}(h).
\end{aligned}$$

The second term on the right behaves like $n^{\frac{5-4\beta}{10}}\phi^{-\frac{9+4\beta}{10}}$. It is now easy to check that, with the choice made for the sequence $\phi(h)$, this defines a convergent series. Also $n\phi(h) \longrightarrow +\infty$, so that the first term also defines a convergent series. The assumption made on β allows the choice of some η such that $\phi(h) \longrightarrow 0$. The convergence of $\widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)$ follows the same arguments with slightly different control on M_γ and the constant c, which is accomplished using **(M1)**. ■

4.2. Asymptotic normality. In this section we prove the asymptotic normality of the random vector (5) following the approach used in the proof of Theorem 1.7 in Bosq [3], based on a large-block small-block decomposition of the sum and a coupling argument for the large-blocks using Lemma 4.2. As for the independent case, the result for the estimator $\widehat{r}_n(x)$ follows using the δ-method.

Let us introduce some notation used for the block decomposition. Let $a, b \in I\!\!R$ and define, for each $n \in I\!\!N$ and $i = 1, \ldots, n$,

$$T_{i,n} = \frac{1}{\sqrt{\phi(h)}}\left[a\left(K\left(\frac{\|x-X_i\|}{h}\right) - I\!\!E K\left(\frac{\|x-X_i\|}{h}\right)\right)\right.$$
$$\left. + b\left(Y_i K\left(\frac{\|x-X_i\|}{h}\right) - I\!\!E\left[Y_i K\left(\frac{\|x-X_i\|}{h}\right)\right]\right)\right].$$

Consider now sequences p_n and q_n of integers such that $p_n + q_n < n$, and define r_n as the largest integer less or equal than $\frac{n}{p_n+q_n}$ (we will drop the explicit reference to the subscript n, for simplicity). We will suppose that $r \sim n^c$, $p \sim n^{1-c}$ and $q \sim n^d$, with $c, d \in (0,1)$ suitably chosen. Let

$$\begin{array}{ll}
V_{1,n} = T_{1,n} + \cdots + T_{p,n}, & V'_{1,n} = T_{p+1,n} + \cdots + T_{p+q,n}, \\
V_{2,n} = T_{p+q+1,n} + \cdots + T_{2p+q,n}, & V'_{2,n} = T_{2p+q+1,n} + \cdots + T_{2(p+q),n}, \\
\vdots & \vdots \\
V_{r,n} = T_{(r-1)(p+q)+1,n} + \cdots + T_{rp+(r-1)q,n}, & V'_{r,n} = T_{rp+(r-1)q+1,n} + \cdots + T_{r(p+q),n}.
\end{array}$$

Finally, define

$$R_n = T_{r(p+q)+1,n} + \cdots + T_{n,n}.$$

Theorem 4.4. *Suppose that* **(K1)**, **(K2)**, **(D1)**, **(D2)**, *with* $\psi(h) \le c_1 \phi^2(h)$, **(R1)**, **(R2)**, **(M1)** *hold. Suppose also that there exists* $a > 1$ *such that*

$$n\phi^a(h) \longrightarrow +\infty, \tag{12}$$

and $(\alpha\mathbf{1})$ *holds with*

$$\beta > \frac{3a}{a-1}. \tag{13}$$

Then, the random vector

$$\sqrt{n\phi(h)}\left(\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x),\, \widehat{g}_n(x) - \mathbb{E}\widehat{g}_n(x)\right)$$

converges in distribution to a centered Gaussian vector with covariance matrix Γ *given by (8).*

Proof: The plan of the proof is as follows: couple the large-blocks $V_{i,n}$ with independent variables, control the distance between the large-blocks and the coupling variables, prove a Lyapounov Theorem for the coupling variables, and prove that the small-blocks and the remaining term are asymptotically negligible.

Step 1. Coupling and controlling the distance. Using recursively Lemma 4.2, with $\theta = 2$, we construct independent random variables $W_{1,n}, \ldots, W_{r,n}$, with distributions $\mathrm{P}_{W_{i,n}} = \mathrm{P}_{V_{i,n}}$, and such that

$$\mathrm{P}\left(|V_{j,n} - W_{j,n}| > u_n\right) \le 11 \left(\frac{\|V_{j,n} + c_n\|_2}{u_n}\right)^{2/5} (\alpha(q))^{4/5}, \tag{14}$$

where $u_n \in (0, \|V_{j,n} + c_n\|_2]$. Now, according to the definition of $T_{i,n}$,

$$\begin{aligned}
&\|T_{i,n}\|_2^2 \\
&= \frac{1}{\phi(h)}\mathrm{Var}\left(aK\left(\frac{\|x - X_i\|}{h}\right) + bYK\left(\frac{\|x - X_i\|}{h}\right)\right) \\
&= \frac{a^2}{\phi(h)}\mathrm{Var}\left(K\left(\frac{\|x - X_i\|}{h}\right)\right) + \frac{b^2}{\phi(h)}\mathrm{Var}\left(Y_iK\left(\frac{\|x - X_i\|}{h}\right)\right) \\
&\qquad + \frac{2ab}{\phi(h)}\,\mathrm{Cov}\left(K\left(\frac{\|x - X_i\|}{h}\right), Y_iK\left(\frac{\|x - X_i\|}{h}\right)\right).
\end{aligned}$$

The first two terms have been shown to converge to $a^2c_2f(x)$ and $b^2g_2(x)f(x)$, respectively. Separating the two mathematical expectations of the covariance in the third term, it is easy to verify that, as for the independent case, it converges to $2abc_2\, r(x)f(x)$. That is

$$\|T_{i,n}\|_2^2 \longrightarrow c_2f(x)\left(a^2 + b^2g_2(x) + 2ab\, r(x)\right) =: B.$$

We apply now the inequality in (b) of Lemma 4.2, with the constant $c_n = 3pB^{1/2}$. With this choice, we have

$$pB^{1/2} \le \|V_{j,n} + c_n\|_2 \le 5pB^{1/2},$$

so that, for n large enough,

$$\mathrm{P}\left(|V_{j,n}-W_{j,n}|>u_n\right)\sim\frac{p^{2/5}(\alpha(q))^{4/5}}{u_n^{2/5}}\,. \tag{15}$$

Finally, at this first step of the proof of the theorem, we look at

$$\Delta_n=\frac{\sum_{j=1}^r W_{j,n}}{\sqrt{rp}}-\frac{\sum_{j=1}^r V_{j,n}}{\sqrt{rp}}\,,$$

which verifies, using (15),

$$\mathrm{P}\left(|\Delta_n|>\varepsilon\right)\leq\sum_{j=1}^r\mathrm{P}\left(|V_{j,n}-W_{j,n}|>\varepsilon\, p^{1/2}\, r^{-1/2}\right)\sim r^{6/5}p^{1/5}(\alpha(q))^{4/5},$$

that is, taking into account the choices made for the block decomposition sequences and $\boldsymbol{(\alpha 1)}$,

$$\mathrm{P}\left(|\Delta_n|>\varepsilon\right)\sim n^{\frac{5c+1-4\beta d}{5}}.$$

This probability converges to zero provided that

$$\beta>\frac{5c+1}{4d}\,. \tag{16}$$

Step 2. The Lyapounov Theorem for the coupling variables. To prove the Central Limit Theorem for $(rp)^{-1/2}\sum_{j=1}^r W_{j,n}$, we will prove that, for some $\rho>2$,

$$z_n=\frac{\sum_{j=1}^r \mathbb{E}(|W_{j,n}|^\rho)}{(r\,\mathrm{Var}(W_{1,n}))^{\rho/2}}\longrightarrow 0. \tag{17}$$

Let us first describe $\mathrm{Var}(W_{1,n})$. It is obvious that $\mathrm{Var}(W_{1,n})=\mathrm{Var}(V_{1,n})$, so we look at

$$\begin{aligned}\frac{1}{p}&\mathrm{Var}(V_{1,n})\\ &=\frac{1}{p\phi(h)}\sum_{i,j=1}^p\mathrm{Cov}\Bigg(aK\left(\frac{\|\,x-X_i\,\|}{h}\right)+bY_iK\left(\frac{\|\,x-X_i\,\|}{h}\right), \qquad (18)\\ &\qquad\qquad aK\left(\frac{\|\,x-X_j\,\|}{h}\right)+bY_jK\left(\frac{\|\,x-X_j\,\|}{h}\right)\Bigg).\end{aligned}$$

The sum is now separated into two sums: the first where $i=j$ and the second one for indexes satisfying $i\neq j$. The term corresponding to the first of these two sums is equal to

$$\frac{1}{\phi(h)}\mathrm{Var}\left(aK\left(\frac{\|x-X\|}{h}\right)+bYK\left(\frac{\|x-X\|}{h}\right)\right),$$

which has been shown to converge to $B=c_2f(x)\Big(a^2+b^2g_2(x)+2ab\,r(x)\Big)$. We now prove that, under our assumptions, the other sum, with the indexes verifying $i\neq j$, converges to zero. As the variables Y_i and the function K are nonnegative valued, it is enough to prove the convergence to zero without the real constants a and b (just replace them by the largest absolute value). To obtain an upper bound we expand the covariance using the bilinearity and proceed by bounding each term. We have then, using **(D2)**, together with the assumption that $\psi(h)\leq c_1\phi^2(h)$, and **(R2)**, where $B(x,h)=\{u\in\mathbf{S}:\ \|x-u\|\leq h\}$,

$$\begin{aligned}
&\mathrm{Cov}\left(K\left(\frac{\|x-X_i\|}{h}\right),K\left(\frac{\|x-X_j\|}{h}\right)\right)\\
&\quad\leq M^2\mathrm{P}(\|x-X_i\|\leq h,\|x-X_i\|\leq h)+M^2\mathrm{P}^2(\|x-X\|\leq h)\\
&\quad= M^2\Big(f_2(x)\psi(h)+f^2(x)\phi^2(h)\Big)\\
&\quad\leq M^2\Big(c_1f_2(x)+f^2(x)\Big)\phi^2(h),
\end{aligned}$$

$$\begin{aligned}
&\mathrm{Cov}\left(Y_iK\left(\frac{\|x-X_i\|}{h}\right),Y_jK\left(\frac{\|x-X_j\|}{h}\right)\right)\\
&\quad\leq M^2\mathbb{E}\Big(s_{ij}(X_i,X_j)I_{B(x,h)\times B(x,h)}(X_i,X_j)\Big)\\
&\qquad\qquad+\left(M\left(\sup_{\|x-u\|\leq h}|r(x)-r(u)|+r(x)\right)f(x)\phi(h)\right)^2\\
&\quad\leq 2M^2s_{ij}(x,x)f_2(x)\psi(h)+4M^2r^2(x)f^2(x)\phi^2(h)\\
&\quad= 2M^2\Big(c_1s_{ij}(x,x)f_2(x)+2r^2(x)f^2(x)\Big)\phi^2(h),
\end{aligned}$$

and finally, using the same kind of arguments,

$$\begin{aligned}
&\mathrm{Cov}\left(K\left(\frac{\|x-X_i\|}{h}\right),Y_jK\left(\frac{\|x-X_j\|}{h}\right)\right)\\
&\qquad\qquad\leq 2M^2\Big(c_1s^*_{ij}(x,x)f_2(x)+r(x)f^2(x)\Big)\phi^2(h).
\end{aligned}$$

So, summarizing, there exists some $t(x)>0$, depending only on $x\in\mathbf{S}$, such that each covariance corresponding to different indexes in (18) is bounded above by $t(x)\phi^2(h)$. Following now the proof of Theorem 1 in Masry [16], we decompose the sum in (18) with distinct indexes into

$$\frac{1}{p\phi(h)}\sum_{\substack{i,j=1\\0<|i-j|\leq a_p}}^{p}\operatorname{Cov}(*,*)+\frac{1}{p\phi(h)}\sum_{\substack{i,j=1\\|i-j|>a_p}}^{p}\operatorname{Cov}(*,*),\tag{19}$$

where we have written $\operatorname{Cov}(*,*)$ for the large covariances appearing in (18), and

$$a_p=\left\lfloor\frac{1}{(\phi(h))^{\frac{1-2/\nu+\eta}{\beta(1-2/\nu)-1}}}\right\rfloor,$$

where $\nu>2$ and $0<\eta<2/\nu$. We apply the upper bound $t(x)\phi^2(h)$ on the first of the sums in (19) to bound it by

$$\frac{2t(x)\phi^2(h)a_p}{\phi(h)}=\frac{2t(x)}{(\phi(h))^{\eta-2/\nu}}\longrightarrow 0.$$

To control the second term, we apply Davidov's inequality (see, for example, Corollary 1.1 in Bosq [3]), to obtain

$$\begin{aligned}
&\frac{1}{p\phi(h)}\sum_{\substack{i,j=1\\|i-j|>a_p}}^{p}\operatorname{Cov}(*,*)\\
&\quad\leq\frac{8t(x)}{p\phi(h)}f^{2/\nu}(x)\phi^{2/\nu}(h)\sum_{\substack{i,j=1\\|i-j|>a_p}}^{p}\big(\alpha(|i-j|)\big)^{1-2/\nu}\\
&\quad\leq\frac{8t(x)f^{2/\nu}(x)}{\phi^{1-2/\nu}(h)}\sum_{l=a_p+1}\big(\alpha(l)\big)^{1-2/\nu}\\
&\quad\sim\frac{a_p^{-\beta(1-2/\nu)+1}}{\phi^{1-2/\nu}(h)}\longrightarrow 0.
\end{aligned}$$

So, summarizing again, we have verified that

$$\operatorname{Var}(W_{1,n})=\operatorname{Var}(V_{1,n})\sim pB=p\,c_2f(x)\Big(a^2+b^2g_2(x)+2ab\,r(x)\Big),$$

thus describing the behaviour of the denominator of (17). We still need to control the numerator of this expression. For this purpose we will apply

Yokoyama's inequality [19], for which we have to control the third moment of the variables $T_{i,n}$. By expanding the power of $T_{i,n}$, we have

$$\begin{aligned}
&\phi^{1/2}(h)\mathbb{E}\left|T_{i,n}\right|^3 \\
&\quad\le \frac{1}{\phi(h)}\mathbb{E}\left[\left(aK\left(\frac{\|x-X_i\|}{h}\right)+bYK\left(\frac{\|x-X_i\|}{h}\right)\right)^3\right] \\
&\qquad+\frac{3}{\phi(h)}\mathbb{E}\left[\left(aK\left(\frac{\|x-X_i\|}{h}\right)+bYK\left(\frac{\|x-X_i\|}{h}\right)\right)^2\right] \\
&\qquad\qquad\times\mathbb{E}\left[aK\left(\frac{\|x-X_i\|}{h}\right)+bYK\left(\frac{\|x-X_i\|}{h}\right)\right] \\
&\qquad+\frac{4}{\phi(h)}\mathbb{E}^3\left[aK\left(\frac{\|x-X_i\|}{h}\right)+bYK\left(\frac{\|x-X_i\|}{h}\right)\right].
\end{aligned}$$

The third term may be rewritten as

$$\phi^2(h)\left(\frac{1}{\phi(h)}\mathbb{E}\left[aK\left(\frac{\|x-X_i\|}{h}\right)+bYK\left(\frac{\|x-X_i\|}{h}\right)\right]\right)^3\longrightarrow 0,$$

as the expression in parenthesis converges to $af(x)+b\,r(x)f(x)$. The second term also converges to zero for analogous reasons. As for the first term, it is less or equal to

$$\begin{aligned}
&\frac{1}{\phi(h)}\left[a^3\mathbb{E}K^3\left(\frac{\|x-X_i\|}{h}\right)+3a^2b\mathbb{E}\left(Y_iK^3\left(\frac{\|x-X_i\|}{h}\right)\right)\right. \\
&\qquad\left.+3ab^2\mathbb{E}\left(Y_i^2K^3\left(\frac{\|x-X_i\|}{h}\right)\right)+b^3\mathbb{E}\left(Y_i^3K^3\left(\frac{\|x-X_i\|}{h}\right)\right)\right].
\end{aligned}$$

Taking absolute values and bounding each K^3 by MK^2, this expression is, up to the multiplication by a constant, bounded by,

$$\begin{aligned}
&\frac{M}{\phi(h)}\left[\mathbb{E}K^2\left(\frac{\|x-X_i\|}{h}\right)+\mathbb{E}\left(Y_iK^2\left(\frac{\|x-X_i\|}{h}\right)\right)\right. \\
&\qquad\left.+\mathbb{E}\left(Y_i^2K^2\left(\frac{\|x-X_i\|}{h}\right)\right)+6M_1^3\mathbb{E}\left(Y_i^2K^2\left(\frac{\|x-X_i\|}{h}\right)\right)\right],
\end{aligned}$$

where we used **(M1)** to bound the final term. Now, we have seen that each of these terms is convergent. That is, we have verified that

$$\phi^{1/2}(h)\mathbb{E}\left|T_{i,n}\right|^3 < \infty, \qquad i = 1, \ldots, n.$$

In order to apply Yokoyama‘s inequality we verify that there exists $\rho \in (2,3)$ such that $\sum_n (n+1)^{\rho/2-1}(\alpha(n))^{(3-\rho)/3} < \infty$. Given the assumptions made on the mixing coefficients, the convergence of this series follows from

$$\beta > \frac{3\,\rho}{2(3-\rho)}\,. \tag{20}$$

Now, applying Yokoyama's inequality it follows that

$$\mathbb{E}\Big((h_n^{\rho/6}\,|W_{j,n}|^{\rho}\Big) \sim p^{\rho/2},$$

so that

$$z_n \sim \frac{r\,p^{\rho/2}}{r^{\rho/2}\,p^{\rho/2}\,h_n^{\rho/6}} = \frac{n^{c-c\rho/2}}{h_n^{\rho/6}}\,, \tag{21}$$

given the choices made for the sequences defining the block sizes. Taking account of (12), this converges to zero provided that $c \geq \frac{\rho}{3\,a(\rho-2)}$. As we want to choose $c \in (0,1)$, we impose that $\rho > \frac{6\,a}{3\,a-1}$. Using this on (20) we need to impose (13). This proves the asymptotic normality of $(rp)^{-1/2}\sum_{j=1}^{r} W_{j,n}$ from which follows the asymptotic normality of $(rp)^{-1/2}\sum_{j=1}^{r} V_{j,n}$. Further, note that it follows from the previous arguments that $\mathrm{Var}(\sum_{j=1}^{r} V_{j,n}) \sim \mathrm{Var}(\sum_{j=1}^{r} W_{j,n}) \sim rp$.

Step 3. Asymptotic negligibility of the remaining terms. Using the same coupling technique it is easy to check that

$$\mathrm{Var}\left(\frac{\sum_{j=1}^{r} V'_{j,n}}{\sqrt{rp}}\right) =\sim \frac{rq}{rp} \sim n^{d-1+c}\,,$$

as the $V'_{j,n}$ are sums of q variables, and

$$\mathrm{Var}\left(\frac{R_n}{\sqrt{rp}}\right) =\sim n^{c-1}\,,$$

for analogous reasons. Thus $(rp)^{-1/2}\,R_n$ converges in probability to zero. The term $(rp)^{-1/2}\sum_{j=1}^{r} V'_{j,n}$ also converges in probability to zero if $d < 1-c$.

Using this on (16), we derive $\beta > (5+a)/(4(a-1))$ which follows from (13), so the theorem is proved. ∎

As the covariance matrix that appears on the limiting distribution is also given by (8), applying the δ-method to the function $\theta(u,v)=\frac{v}{u}$ proves the same result as in the previous section.

Theorem 4.5. *Suppose all the assumptions of Theorem 4.4 are satisfied. Then*

$$\sqrt{n\phi(h)}\Big(\widehat{r}_n(x)-\mathbb{E}\widehat{r}_n(x)\Big)\xrightarrow{d}\mathcal{N}\left(0,c_2\left(g_2(x)-\frac{r^2(x)}{f(x)}\right)\right).$$

References

[1] Bensaïd, N. Fabre, J.-P., Convergence de l'estimateur à noyau de dérivées de Radon-Nikodym générales dans le cas mélangeant, Canad. J. Statist. 26, 267–282, 1998.

[2] Bensaïd, N., Oliveira, P.E., Histogram estimation of Radon-Nikodym derivatives for strong mixing data, Statistics 35, 569–592, 2001.

[3] Bosq, D., *Nonparametric statistics for stochastic processes*, Lect. Notes in Statist. 110, Springer, 1996.

[4] Bradley, R., Approximation theorems for strongly mixing random variables, Michigan Math. J. 30, 69–81, 1983.

[5] Dabo-Niang, S., Kernel density estimator in an infinite-dimensional space with a rate of convergence in the case of diffusion process, Appl. Math. Lett. 17, 381–386, 2004.

[6] Ferraty, F., Vieu, P., Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés, C. R. Acad. Sci. Paris Sér. I Math. 330, 139–142, 2000.

[7] Ferraty, F., Vieu, P., Nonparametric models for functional data, with application in regression, time series prediction and curve estimation, The International Conference on Recent Trends and Directions in Nonparametric Statistics, J. Nonparametr. Stat. 16, 111–125, 2004.

[8] Ferrieux, D., Estimation of the mean measure density of a discrete random measure through associated sequences of observations, Statistics 33, 129–152, 1999.

[9] Gasser, T, Hall, P., Presnell, B., Nonparametric estimation of the mode of a distribution of random curves, J. Roy. Statist. Soc. Ser. B 60, 681–691, 1998.

[10] Geffroy, J., Sur l'estimation d'une densité dans un espace métrique, C. R. Acad. Sci. Paris Sér. A 278, 1449–1452, 1974.

[11] Geffroy, J., Etude de la convergence du régressogramme, Pub. Inst. Statist. Univ. Paris 25, 41–56, 1980.

[12] Jacob P., Niéré, L., Contribution à l'estimation des lois de Palm d'une mesure aléatoire, Pub. Inst. Stat. Univ. Paris 35, 39–49, 1990.

[13] Jacob, P., Oliveira, P.E., A general approach to nonparametric histogram estimation, Statistics 27, 73–92, 1995.

[14] Jacob, P., Oliveira, P.E., Kernel estimators of general Radon-Nikodym derivatives, Statistics 30, 25–46, 1997.

[15] Jacob, P., Oliveira, P.E., Histograms and associated point processes, Stat. Inference Stoch. Process. 2, 227–251, 1999.

[16] Masry, E., Nonparametric regression estimation for dependent functional data: asymptotic normality, Stochastic Process. Appl. 115, 155–177, 2005.

[17] Ramsay, J., Silverman, B., Applied Functional Data Analysis: Methods and Case Studies, Springer, New York, 2002.
[18] Rhomari, N., *Filtrage non paramétrique pour les processus non Markoviens. Applications*, Ph.D. Dissertation, Univ. Pierre et Marie Curie, Paris, 1994.
[19] Yokoyama, R., Moment bounds for stationary absolutely regular sequences, Z. Wahrsch. Verw. Gebiete 52, 45–57, 1980.

Paulo Eduardo Oliveira
Dep. Matemática, Univ. Coimbra, Apartado 3008, 3001 - 454 Coimbra, Portugal