Journal of
Mathematical
Inequalities
Volume 17, Number 2 (2023), 481–515 doi:10.7153/jmi-2023-17-32

ASYMPTOTIC DISTRIBUTION OF THE WAVELET–BASED ESTIMATORS OF MULTIVARIATE REGRESSION FUNCTIONS UNDER WEAK DEPENDENCE

SOUMAYA ALLAOUI*, SALIM BOUZEBDA AND JICHENG LIU

(*Communicated by M. Krnić*)

Abstract. This paper investigates the nonparametric linear wavelet-based estimators of multivariate regression functions. Under mild conditions, we establish the asymptotic normality under the weak dependence, which incorporates mixing and association concepts. This framework applies to numerous classes of intriguing statistical processes, primarily Gaussian sequences and, more generally, Bernoulli shifts. We give an application for the confidence interval.

1. Introduction

The regression analysis has demonstrated its adaptability and provided a robust statistical modeling framework in various applied and theoretical contexts where the predictive relationship between related responses and predictors is to be modeled. It is important to note that parametric regression models provide useful tools for analyzing practical data when the models are correctly specified but may be susceptible to large modeling biases if the model structures are incorrectly specified, which is the case for many practical problems. As an alternative, nonparametric smoothing techniques alleviate modeling bias concerns. In this article, we will focus on the study of estimators of the wavelet type. Let $(\mathbf{X},Y)$ be $\mathbb{R}^d\times\mathbb{R}$ valued random variable with common joint Lebesgue density $f_{\mathbf{X}Y}(\cdot)$ and marginals $f_{\mathbf{X}}(\cdot)$ and $f_Y(\cdot)$. For a chosen measurable function $\varphi(\cdot)$ and $\mathbf{x}\in\mathbb{R}^d$, the regression function, whenever it exists, is defined to be

$$m(\varphi,\mathbf{x}) := \mathbb{E}(\varphi(Y)\mid\mathbf{X}=\mathbf{x}). \tag{1.1}$$

A well-known estimator for the regression function $m(\cdot,\varphi)$, often used in nonparametric statistics, is the kernel regression function estimator. This estimator is, under suitable conditions, strongly consistent, i.e., it converges almost surely to the unknown regression function at $\mathbf{x}$. Similarly to the kernel density estimator of $f_{\mathbf{X}}(\cdot)$, which is, under suitable conditions, strongly consistent. Because of numerous applications and their important role in mathematical statistics, the problem of estimating $m(\cdot,\varphi)$ and

Mathematics subject classification (2020): 62E20, 60F05, 62G08, 62H12, 62G09.

Keywords and phrases: Multivariate regression estimation, wavelets basis, Besov spaces, weakly dependent processes, stationarity, central limit theorem, confidence intervals.

* Corresponding author.

$f_{\mathbf{X}}(\cdot)$ has been a subject of considerable interest during the last decades. For good sources of references to the research literature in this area along with statistical applications, consult [20], [51], [43], [53], [28], [21], [32], [50], [14], and the references therein. However, a major problem with this approach appears in the estimation of compactly supported or discontinuous curves at boundary points. In the kernel method estimation, we typically require that the density satisfy certain smoothness conditions like two-times continuous differentiability. In situations where the density does not fulfill these requirements, the wavelet method is an attractive alternative that often performs relatively well by the fact that it adapts automatically to the regularity of the curve to be estimated. Wavelet estimators assume that the underlying curve belongs to a function space with certain degrees of smoothness. The wavelet estimators do not depend on the smoothness parameters; nevertheless, they behave as if the true curve is known in advance and attain the optimal convergence rates. In the present contribution, the smoothness assumptions on the underlying regression curve are considerably relaxed by using the wavelets methods in estimating functions in Besov spaces. For more details on wavelet's theory, we refer to [41], [17] and [52], among others. The statistical curve estimation using wavelets methodology is surveyed in [35]. For wavelet linear estimators in various settings, we can refer to ([37, 38, 39]), [57], [15, 16]. However, the asymptotic normality of wavelet estimators has received little attention. [5] have established strong consistency in the supremum norm, and the asymptotic normality results for the wavelet density and regression function estimators in the setting of the $\mathbb{R}^d$-valued ergodic process. These results were extended to time continuous $\mathbb{R}^d$-valued ergodic process by [12], for recent references see [7, 6], [23, 22].

Our aim in this work is to relax the concept of mixing conditions that are quite restrictive or even fail to fit some processes of interest; for instance, take an AR(1)-input solution of the recursion

$$X_k = \frac{1}{b}\left(X_{k-1} + \xi_k\right), \quad k \in \mathbb{Z},$$

where $b \geqslant 2$ is an integer and $(\xi_k)_{k\in\mathbb{N}}$ are independent and uniformly distributed random variables on the set $U(b) := \{0, 1, \ldots, b-1\}$. This process is not mixing in the sense of Rosenblatt, as this is shown in [3] for $b = 2$; however, [26] proved that such a process is weakly dependent. To overcome this situation, some new ideas of weak dependence proposed by [26] and [4] leading to the dependence structure in terms of covariance bounds of functions defined on Lipschitz bounded or bounded variations spaces rather than uniformly bounded functions ones and sigma algebras as in mixing. In the last reference, another type of weak dependence, was introduced. It turns out that the notion of weak dependence is more general than mixing and allows for treatment; for example, all the usual causal or non-causal time series are weakly dependent processes: this is the case, for instance, of Gaussian, associated, linear, ARCH(∞), etc. Inference techniques for weakly dependent processes have gained importance because of their relevance in modern applications. In this framework, only a few results on central limit theorems have been investigated, refer to [27]. [19] in Section 2.2.3, proposed dependence coefficients as distances of conditional expectations of quadrants indicators such the supremum is taken over a sequence of σ-algebras from the past and several points of the sequence of the variables defined on $\mathbb{R}^{dn}$ in the future. We consider in the present paper the weakest coefficient type of $\widetilde{\alpha}(r)$, defined bellow, that is appropriate for the study of the time series problems.

To the best of our knowledge, the results presented here respond to a problem that has not been studied systematically until recently, giving this paper the primary motivation. Indeed, our paper is to provide a first full theoretical justification of the normality of the wavelet estimation for multivariate regression functions when the observed data are assumed to be generated from an $\mathbb{R}^d-$ process under weak dependence. The present paper complements our previous work [2], where we have provided strong uniform consistency properties with rates of the linear wavelet estimators over compact subsets of $\mathbb{R}^d$.

The rest of the paper is structured as follows. In Section 2, we recall some basic definitions for wavelets and Besov spaces. In Section 3, we give some details concerning the weak dependence, Subsection 3.1, and introduce the linear wavelet estimators of density and regression functions, Subsection 3.2. In Section 4, we give the assumptions with some comments and asymptotic normality of the linear wavelet estimators together with an application for the confidence intervals. We conclude our paper with some concluding remarks and possible future developments in Section 5. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to Section 6.

2. Wavelets and Besov space

In this section, we briefly introduce notation corresponding to wavelets and Besov spaces. First, we provide a general introduction to wavelet multiresolution theory, which is detailed in [41] and [37, 38, 39]. Define $\{V_j\}_{j=1}^{\infty}$ a multiresolution analysis of $\mathbf{L}^2(\mathbb{R}^d)$ [1] as a decomposition of the space $\mathbf{L}^2(\mathbb{R}^d)$ into an increasing sequence of closed subspaces $\{V_j : j \in \mathbb{Z}\}$ such that:

(i) $V_j \subset V_{j+1},\ \ j \in \mathbb{Z}$, where $\mathbb{Z}$ denotes the set of integers,

(ii) $\cap_j V_j = \emptyset,\ \ \overline{\cup_j V_j} = \mathbf{L}^2(\mathbb{R}^d)$,

(iii) $f(\mathbf{x}) \in V_j \Leftrightarrow f(2\mathbf{x}) \in V_{j+1},\ \ f(\mathbf{x}) \in V_j \Rightarrow f(\mathbf{x}+\mathbf{k}) \in V_j,\ \ \mathbf{k} \in \mathbb{Z}^d$,

(iv) There exists a scaling function $\phi(\cdot) \in \mathbf{L}^2(\mathbb{R}^d)$, integrated to 1 on $\mathbb{R}^d$ such that

$$\left\{\phi_k(\mathbf{x}) = \phi\left(\mathbf{x}-\mathbf{k}\right), \mathbf{k} \in \mathbb{Z}^d\right\}$$

forms an orthonormal basis on $V_0 = \overline{span}\{\phi(\cdot - \mathbf{k})\}$ of $\mathbf{L}^2\left(\mathbb{R}^d\right)$,

[1] The vector space $L_p(\mathbb{R}^d), 1 \leqslant p < \infty$ is the set of all measurable functions such that $\int_{\mathbb{R}^d} |f(x)|^p dx < \infty$. The norm is defined by

$$\|f\|_{\mathbf{L}_p} = \left(\int_{\mathbb{R}^d} |f(\mathbf{x})|^p d\mathbf{x}\right)^{1/p}.$$

by that,

$$\{\phi_{j,\mathbf{k}}(\mathbf{x}) = 2^{jd/2}\phi(2^j\mathbf{x}-\mathbf{k}) : \mathbf{k} \in \mathbb{Z}^d\}$$

constitute an orthonormal basis for V_j. For the purpose of the paper, $\phi(\cdot)$ is supposed to be r-regular ($r \geqslant 1$), i.e., the scaling function $\phi(\cdot)$ belongs to the class $\mathscr{C}^r$ and satisfies, for each $i > 0$,

$$\left|\left(D^{\boldsymbol{\beta}}\phi\right)(\mathbf{x})\right| \leqslant \frac{C_i}{(1+\|\mathbf{x}\|)^i}, \text{ for all } |\boldsymbol{\beta}| \leqslant r, \tag{2.1}$$

where C_i is constant depending only on i,

$$\left(D^{\boldsymbol{\beta}}\phi\right)(\mathbf{x}) = \frac{\partial^{\boldsymbol{\beta}}\phi(\mathbf{x})}{\partial^{\beta_1}x_1\cdots\partial^{\beta_d}x_d}$$

and

$$\boldsymbol{\beta} = (\beta_1,\ldots,\beta_d), \quad |\boldsymbol{\beta}| = \sum_{i=1}^{d}\beta_i.$$

Take

$$V_j \oplus W_j = V_{j+1}.$$

Following the methodology of [17], one may derive $N = 2^d - 1$ associated wavelet functions $\{\psi_i; i = 1,\ldots,N\}$ such that

(W.1) $\{\psi_i(\mathbf{x}-\mathbf{k}) : \mathbf{k} \in \mathbb{Z}^d : i = 1,\ldots,N\}$ is an orthonormal basis for W_0,

(W.2) the family $\{\psi_{i,j,\mathbf{k}}(\mathbf{x}) : i = 1,\ldots,N, \mathbf{k} \in \mathbb{Z}^d, j \in \mathbb{Z}\}$ composed by functions

$$\psi_{i,j,\mathbf{k}}(\mathbf{x}) = 2^{jd/2}\psi_i(2^j\mathbf{x}-\mathbf{k})$$

forms an orthonormal basis for $\mathbf{L}^2(\mathbb{R}^d)$,

(W.3) $\psi_i(\cdot)$ has the same regularity as $\phi(\cdot)$ and both functions have compact support $[-L,L]^d$ for some $L > 0$.

Then, starting from an integer j_0, any $f_{\mathbf{X}}(\cdot) \in \mathbf{L}^2(\mathbb{R}^d)$ can be represented as

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} a_{j_0,\mathbf{k}}\phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{i=1}^{2^d-1}\sum_{j=j_0}^{\infty}\sum_{\mathbf{k}\in\mathbb{Z}^d} b_{i,j,\mathbf{k}}\psi_{i,j,\mathbf{k}}(\mathbf{x}) \tag{2.2}$$

where, $a_{j_0,\mathbf{k}}$ and $b_{i,j,\mathbf{k}}$ represent the wavelet coefficients given as

$$a_{j_0,\mathbf{k}} = \int_{\mathbb{R}^d} f(\mathbf{u})\phi_{j_0,\mathbf{k}}(\mathbf{u})d\mathbf{u}$$

and

$$b_{i,j,\mathbf{k}} = \int_{\mathbb{R}^d} f(\mathbf{u})\psi_{i,j,\mathbf{k}}(\mathbf{u})d\mathbf{u}.$$

The orthogonal projection of $f_{\mathbf{X}}(\cdot)$ on some subspace $V_\ell \subset \mathbf{L}^2(\mathbb{R}^d)$ can be written in two equivalent ways, for any $j_0 \leqslant \ell$,

$$\begin{aligned}\left(P_{V_\ell} f\right)(\mathbf{x}) &:= \sum_{\mathbf{k}\in\mathbb{Z}^d} a_{\ell,\mathbf{k}}\phi_{\ell,\mathbf{k}}(\mathbf{x}) \\ &= \sum_{\mathbf{k}\in\mathbb{Z}^d} a_{j_0,\mathbf{k}}\phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{j=j_0}^{\ell}\sum_{i=1}^{N}\sum_{\mathbf{k}\in\mathbb{Z}^d} b_{i,j,\mathbf{k}}\psi_{i,j,\mathbf{k}}(\mathbf{x}). \end{aligned} \tag{2.3}$$

As usual in the wavelet estimation literature, the Besov spaces $\mathbf{B}_{s,p,q}$ are considered essential for their exceptional expressive power in describing the smoothness proprieties in functional estimation and approximation theory they contain a wide variety of homogeneous and inhomogeneous functions spaces used in statistical research. In terms of wavelet coefficients, [41] characterized the Besov space as follows: for $0 < s < r$ is a real-valued smoothness parameter of $f \in \mathbf{L}^p(\mathbb{R}^d)$, then $f \in \mathbf{B}_{s,p,q}$ equivalent to

(B.1) $J_{s,p,q}(f) = \|P_{V_0}f\|_{\mathbf{L}^p} + \left(\sum_{j\geqslant 0}\left(2^{js}\|P_{W_j}f\|_{\mathbf{L}^p}\right)^q\right)^{1/q} < \infty,$

(B.2) $J'_{s,p,q}(f) = \|a_0\cdot\|_{l_p} + \left(\sum_{j\geqslant 0}\left(2^{j(s+d(1/2-1/p))}\|b_j\cdot\|_{l_p}\right)^q\right)^{1/q} < \infty$

with

$$\|a_0\cdot\|_{l_p} = \left(\sum_{\mathbf{k}\in\mathbb{Z}^d}|a_{0,\mathbf{k}}|^p\right)^{1/p}$$

and

$$\|b_j\cdot\|_{l_p} = \left(\sum_{i=1}^{N}\sum_{\mathbf{k}\in\mathbb{Z}^d}|b_{i,j,\mathbf{k}}|^p\right)^{1/p}$$

and with the usual sup-norm modification for $q=\infty$. For more equivalent characterizations of Besov space and its advantages in approximation theory and statistics, we refer to [24].

3. Statistical framework

Now we give the notation and definitions needed for the detailed statement of our results concerning the setting of the weak dependence conditions and the wavelet estimator of the multivariate regression function. Let $\mathfrak{C}$ and M be some positive constants that may differ from one term to the other. $\|X\|_p := [\mathbb{E}(|X|^p)]^{1/p}$, and $\mathbb{1}_A(\cdot)$ is the indicator function of A. Also for the numerical sequences of positive constants a_n, b_n, where $\{a_n, n \geqslant 1\}$ and $\{b_n, n\geqslant 1\}$, we have the following notation: $a_n = o(b_n)$, (The sequence a_n is negligible with respect to the sequence b_n) if for all $\varepsilon > 0$, there is some integer $n_0 \in \mathbb{N}$ such that $a_n \leqslant \varepsilon b_n$ for all integers $n \geqslant n_0$. $a_n = O(b_n)$ (The sequence a_n is dominated by the sequence b_n), if for a relevant constant $C > 0$, $a_n \leqslant C b_n$ for all integers $n\in\mathbb{N}$. Similarly, the sequences a_n and b_n have the same order, i.e., $a_n \asymp b_n$ if $a_n = O(b_n)$ and $b_n = O(a_n)$. Throughout the paper, any mathematical symbol referring to a multivariate point $\mathbf{x} = (x_1,\ldots,x_{d_{\mathbf{x}}}) \in \mathbb{R}^{d_{\mathbf{x}}}$ will be print in bold symbol.

3.1. Weak dependence notion

The idea of general notion of weak dependence can be summarized in the following way. Consider two finite samples with time indices Pa in the past and the future Fu, separated by a gap r. The independence of Pa and Fu is equivalent to $\operatorname{cov}(g_1(Fu), g_2(Pa)) = 0$ for a suitable class of measurable functions. A natural way to weaken this condition is to provide precise control of these covariances as the gap r becomes larger, and to fix the rate of decrease of the control as r tends to infinity. We will precise in this section the dependence that we will consider. Let us introduce

$$\mathbf{L}^\infty = \bigcup_{n=1}^{\infty} \mathbf{L}^\infty(\mathbb{R}^n),$$

the set of real-valued and bounded functions on the space $\mathbb{R}^n$ for $n = 1, 2, \ldots$. Consider a function $g : \mathbb{R}^n \to \mathbb{R}$ where $\mathbb{R}^n$ is equipped with its ℓ^1-norm (i.e., $\|(x_1, \ldots, x_n)\|_1 = |x_1| + \cdots + |x_n|$) and define the Lipschitz modulus of $g(\cdot)$ as

$$\operatorname{Lip}(g) = \sup_{\mathbf{x} \neq \mathbf{y}} \frac{|g(\mathbf{x}) - g(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_1}.$$

This paper deals with the estimation of the regression function by wavelet estimators for a strictly stationary d-dimensional weakly dependent process. This investigation is not reported in previous statistical frameworks. Besides, as was mentioned before, in our contribution we give assumptions of $\tilde{\alpha}$-weakly dependence on the process. First, we recall the definition of $\widetilde{\alpha}(r)$ coefficient from [19, Definition 2.5].

DEFINITION 3.1. Let $\mathbf{X}$ be an $\mathbb{R}^d$-valued random process defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ an n copies of $\mathbf{X}$. For $\mathbf{t} = (\mathbf{t}_1, \ldots, \mathbf{t}_n)$ a n-tuple of $\mathbf{t}_i \in \mathbb{R}^d$ and for $\mathbf{x} \in \mathbb{R}^d$, define

$$g_{\mathbf{t},i}(\mathbf{x}) = \mathbb{1}_{\{\mathbf{x} \leqslant \mathbf{t}_i\}} - \mathbb{P}(\mathbf{X}_i \leqslant \mathbf{t}_i).$$

The coefficient $\widetilde{\alpha}(\mathscr{M}, \mathbf{X})$ is given by the equation:

$$\widetilde{\alpha}(\mathscr{M}, \mathbf{X}) = \sup_{\mathbf{t} \in \mathbb{R}^{dn}} \left\| \int \prod_{i=1}^{n} g_{\mathbf{t},i}(x_i) \mathbb{P}_{\mathbf{X}|\mathscr{M}}(d\mathbf{x}) - \mathbb{E} \prod_{i=1}^{n} g_{\mathbf{t},i}(\mathbf{X}_i) \right\|_1,$$

where $\mathbb{P}_{\mathbf{X}|\mathscr{M}}$ is the conditional distribution of $\mathbf{X}$ given $\mathscr{M}$ a sub-sigma algebra of $\mathscr{F}$.

Take $(d_1, d_2) \in \mathbb{N}^{*2}$ and write $\Delta(d_1, d_2, r)$ be the set of $(i, j) \in \mathbb{Z}^{d_1} \times \mathbb{Z}^{d_2}$ such that

$$\Delta(d_1, d_2, r) = i_1 \leqslant \cdots \leqslant i_{d_1} < i_{d_1} + r \leqslant j_1 \leqslant \cdots \leqslant j_{d_2},$$

where $r = |j_1 - i_{d_1}| \geqslant 0$. Let $(\mathscr{M}_i)_{i \in \mathbb{Z}}$ be the sequence of sub-σ-algebras of $\mathscr{F}$. For the general case when $(\mathbf{X}_i)_{i \in \mathbb{Z}}$ is a sequence of $\mathbb{R}^d$ valued random variables, we define

$$\widetilde{\alpha}_n(r) = \max_{1 \leqslant v \leqslant n} \sup_{(i,j) \in \Delta(1, d_2, r)} \widetilde{\alpha}\left(\mathscr{M}_i, \left(\mathbf{X}_{j_1}, \ldots, \mathbf{X}_{j_{d_2}}\right)\right).$$

The main advantage of this definition is that it allows measuring the dependence between the sequences $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ and $(\mathscr{M}_i)_{i\in\mathbb{Z}}$ by considering d_2-tuples of some process $\mathbf{X}$ in the future. According to [19, Remark 2.4, p17]), in the special case of $n=1$, $d=1$, i.e., X is an $\mathbb{R}$-valued random variable, the coefficient $\widetilde{\alpha}(\mathscr{M},X)$ was introduced by [47] as follows

$$\widetilde{\alpha}(\mathscr{M},X)=\sup_{x\in\mathbb{R}}\mathbb{E}|\mathbb{P}(X\leqslant x|\mathscr{M})-\mathbb{P}(X\leqslant x)|.$$

Notice that the usual $\alpha(\mathscr{M},\sigma(X))$-mixing coefficient defined by [48] satisfies

$$\widetilde{\alpha}(\mathscr{M},X)\leqslant\alpha(\mathscr{M},\sigma(X)).$$

An $\mathbb{R}^d$-process $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ is said to be $\widetilde{\alpha}$-weak dependent if

$$\widetilde{\alpha}(r)=\sup_{n\in\mathbb{N}}\widetilde{\alpha}_n(r), \tag{3.1}$$

tends to 0 as whenever the gap r between indices of the initial time series in the past and the future terms goes to infinity.

PROPOSITION 3.1. [47, Eq (1.11c)] *Let X,Y be a $\mathbb{R}$-valued random variable with $\mathbb{E}(Y)=0$. Then, we have*

$$|\mathrm{Cov}\,(X,Y)|\leqslant 2\int_0^{\alpha(\mathscr{M},Y)}Q_X(u)Q_Y(u)du\leqslant 2\alpha(\mathscr{M},Y)^{1/p}\|X\|_q\|Y\|_s,$$

where $Q_{|X|}$ is the generalized inverse of X defined for any $u\in[0,1]$ by

$$Q_X(u)=\inf\{x\in\mathbb{R}:\mathbb{P}(|X|>x)\leqslant u\}$$

and p,q and s are strictly positive reals satisfy $p^{-1}+q^{-1}+s^{-1}=1$. If X and Y are almost surely bounded, the upper bound holds for $p=1,(q,s=\infty)$.

In the following proposition, we extract a covariance upper bound involving the coefficient of $\widetilde{\alpha}$ that will be used in the sequel.

PROPOSITION 3.2. *Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space.*

Denote $\mathbf{X}_{i,d_1}:=(X_{i_1},\ldots,X_{i_{d_1}})$ an $\mathbb{R}^{d_1}$ random vector adapted $\mathscr{M}$ a σ-algebra of $\mathscr{F}$. Note $\mathbf{X}_{j,d_2}:=(X_{j_1},\ldots,X_{j_{d_2}})$ an $\mathbb{R}^{d_2}$-valued process distributed as $\mathbf{X}_i$ and independent of $\mathscr{M}$, where $|j_1-i_{d_1}|>0$ and satisfies $\|\mathbf{X}_j\|_2<\infty$. If Ω is rich enough, for a bounded function $g_1(\cdot)$, and Lipschitz function $g_2(\cdot)$, the $\widetilde{\alpha}$-dependence coefficient provides the covariance bounds as

$$\left|\mathrm{cov}\left(g_1\left(\mathbf{X}_{i,d_1}\right),g_2\left(\mathbf{X}_{j,d_2}\right)\right)\right|\leqslant C\|g_1\left(\mathbf{X}_i\right)\|_\infty\mathrm{Lip}(g_2)d_2\widetilde{\alpha}^{\frac{1}{2}}(r), \tag{3.2}$$

where $C=4\,\|\mathbf{X}\|_2$.

The proof of Proposition 3.2 is given in Section 7.

REMARK 3.1. The last relation shows that a property of weak dependence follows in the sense that the covariance of functions belonging to some regular spaces[2] tends to zero as the time gap between the two blocks of observations increases. The $\tilde{\alpha}$-weakly dependent in this setting satisfies the dependence notion of [26] [3] for

$$\Psi(d_1,d_2,g_1,g_2) = C\|g_1\|_\infty \mathrm{Lip}(g_2)d_2.$$

By choosing the dependence coefficient

$$\theta(r) = \tilde{\alpha}^{\frac{1}{2}}(r).$$

This includes the θ-weak dependence (see [19]) which is the causal counterpart of η coefficients defined in [27]. The Proposition 3.2 assumptions can be verified for a large class of Markov chains.

3.1.1. Examples of weakly dependent processes

The general concept of weak dependence has been shown to efficiently treat the dependence structure of large classes of non-mixing models. To conclude this section we give examples of Bernoulli shifts and Markov chains models for which the concept of $\tilde{\alpha}$-dependence holds, we refer to [25] and [19, Chapter 3], for detailed examples. Let us introduce the following notions that will be used in the sequel. For some $p \in [0,\infty]$, we define a non increasing sequence $\delta_{i,p}$ such

$$\|X_i - X_i^*\|_p \leqslant \delta_{i,p}, \tag{3.3}$$

or

$$(\mathbb{E}\|X_i - X_i^*\|^p)^{\frac{1}{p}} \leqslant \delta_{i,p}, \tag{3.4}$$

for X_i^* is a random variable (or vector) distributed as X_i and independent of $\mathscr{M}_0$, the σ-algebra generated by X_i.

3.1.1.a. Causal Bernoulli shifts

Let μ be a probability distribution on a measurable space $(E,\mathscr{E})$. Consider an i.i.d. sequence $(\xi_n)_{n\in\mathbb{Z}}$ with marginal law μ. Let $\nu = \mu^{\otimes\mathbb{Z}}$ be the law of $(\xi_n)_{n\in\mathbb{Z}}$

[2]The space of real-valued and Lipschitz functions.

[3] Let $\mathscr{F}$ be a class of real-valued functions, such that for each $L(\cdot) \in \mathscr{F}$ there exists an integer $n \geqslant 1$ such that $L(\cdot)$ is defined on $\mathbb{R}^n$.

DEFINITION 3.2. ([26]) The sequence $(X_n)_{n\in\mathbb{N}}$ of r.v.s is called $(\theta,\mathscr{F},\psi)$-weak dependent, if there exists a class $\mathscr{F}$ of real-valued functions, a sequence $\theta = (\theta(r))_{r\in\mathbb{N}}$ decreasing to zero at infinity, and a function ψ with arguments $(h,k,d_1,d_2) \in \mathscr{F}^2 \times \mathbb{N}^2$ such that for any d_1-tuple $(i_1,\ldots,i_{d_1})$ and any d_2-tuple $(j_1,\ldots,j_{d_2})$ with $i_1 \leqslant \cdots \leqslant i_u < i_{d_1} + r \leqslant j_1 \leqslant \cdots \leqslant j_{d_2}$, one has

$$\left|\mathrm{Cov}\left(h\left(X_{i_1},\ldots,X_{i_{d_1}}\right),k\left(X_{j_1},\ldots,X_{j_{d_2}}\right)\right)\right| \leqslant \psi(h,k,d_1,d_2)\theta(r),$$

for all functions $h,k \in \mathscr{F}$ that are defined respectively on $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$.

on the space $\left(E^{\mathbb{Z}}, \mathscr{E}^{\otimes \mathbb{Z}}\right)$. Then $\mathbf{L}^p(E^{\mathbb{Z}})$ is the space of measurable functions ν -a.s defined on $E^{\mathbb{Z}}$ and such that

$$\mathbb{E}\left|H\left((\xi_n)_{n\in\mathbb{Z}}\right)\right|^p < \infty.$$

Analogously, we let $\nu^+ = \mu^{\otimes\mathbb{N}}$ for the law of $(\xi_n)_{n\in\mathbb{N}}$ on the space $\left(E^{\mathbb{N}}, \mathscr{E}^{\otimes\mathbb{N}}\right)$ A Bernoulli shift is an $\mathbf{L}^p$-stationary process defined as

$$X_i = H\left((\xi_{i-j})_{j\in\mathbb{Z}}\right), \qquad \text{for some} \quad H \in \mathbf{L}^p(E^{\mathbb{N}}).$$

A causal Bernoulli shift is associated with $H \in \mathbf{L}^p\left(E^{\mathbb{N}}\right)$. To obtain a bound for the coefficients, we introduce some regularity conditions on function $H(\cdot)$. In the following, we consider two special cases of causal Bernoulli shifts. When the innovations ξ_j are assumed to be independent and identically distributed, we refer to the causal shifts with independent inputs. Suppose that (3.3) satisfied with $X_i^* = H\left(\xi^*_{i-j}, j\in\mathbb{Z}\right)$, such $\xi_j^* = \xi_j$ if $j > 0$ and $\xi_j^* = \xi_j'$ for $j \leqslant 0$ for an independent copy $\left(\xi_j'\right)_{j\in\mathbb{Z}}$ of $(\xi_j)_{j\in\mathbb{Z}}$, and $\mathscr{M}_0 = \sigma(X_j, j \leqslant 0)$. If X_0 has continuous distribution function, with modulus of uniform continuity $\omega(\cdot)$, then the Bernoulli shifts X_i are $\tilde{\alpha}$-dependent with

$$\tilde{\alpha}_i(r) \leqslant 2r\left(\frac{\delta_{p,i}}{g_p^{-1}(\delta_{p,i})}\right)p,$$

where $g_p(y) = y(\omega(y))^{1/p}$, for $p \in [0,\infty[$. Furthermore Assume that X_0 has a continuous distribution function, with a modulus of uniform continuity $\omega(\cdot)$. Then, we have

$$\tilde{\alpha}_i(r) \leqslant r\omega(\delta_{\infty,i}).$$

The same results are verified, for causal shifts with dependent inputs, by taking the bound (3.4).

3.1.1.b. Markov sequences

Let $(X_n)_{n\geqslant 1-d}$ be a sequence of random variables with values in a Banach space $(\mathbb{R}, \|\cdot\|)$. Assume that X_n satisfies the recurrence equation

$$X_n = H(X_{n-1}, \ldots, X_{n-d}; \xi_n), \tag{3.5}$$

where $H(\cdot)$ is a measurable function with values in $\mathbb{B}$, the sequence $(\xi_n)_{n>0}$ is i.i.d. and $(\xi_n)_{n>0}$ is independent of $(X_0, \ldots, X_{d-1})$. Note that if X_n satisfies (3.5) then the random variable $\mathbf{Y}_n = (X_n, \ldots, X_{n-d+1})$ defines a Markov chain such that $\mathbf{Y}_n = M(\mathbf{Y}_{n-1}; \xi_n)$ with

$$M(x_1, \ldots, x_d; \xi) = (H(x_1, \ldots, x_d; \xi), x_1, \ldots, x_{d-1}).$$

Assume that $(X_n)_{n\leqslant d-1}$ is a stationary solution to (3.5). Let $Y_0 = (X_0, \ldots, X_{1-d})$, and let $\mathbf{Y}_0^* = \left(X_0^*, \ldots, X_{1-d}^*\right)$ be and independent vectors with the same law as $\mathbf{Y}_0$ (that is a distribution invariant by $M(\cdot)$). Let then $X_n^* = F\left(X_{n-1}^*, \ldots, X_{n-d}^*; \xi_n\right)$. Clearly, for $n > 0$, X_n^* is distributed as X_n and independent of $\mathscr{M}_0 = \sigma\left(X_i, 1-d \leqslant i \leqslant 0\right)$. Assume the bound (3.4), then the same results for Bernoulli shift are verified with same assumptions on distribution functions. Suppose

$$\left(\mathbb{E}\left\|H\left(x;\xi_1\right)-H\left(y;\xi_1\right)\right\|^p\right)^{\frac{1}{p}} \leqslant \sum_{i=1}^d a_i\left\|x_i-y_i\right\|, \quad \sum_{i=1}^d a_i < 1. \tag{3.6}$$

If (3.6) is verified for $p=1$ and the distribution function $F_X(\cdot)$ of X_0 satisfies

$$|F_X(x)-F_X(y)| \leqslant K|x-y|^\gamma, \text{ for } \gamma \in]0,1],$$

then,for $\rho \in [0,1[$, we have the upper bound

$$\tilde{\alpha}(r) \leqslant 2nK^{1/(\gamma+1)}C^{\gamma/(\gamma+1)}\rho^{r\gamma/(\gamma+1)}.$$

If the condition (3.6) holds for $p=\infty$, then

$$\tilde{\alpha}(r) \leqslant nKC^\alpha\rho^{\gamma r}.$$

3.2. Linear wavelets regression estimators

Let $\{\mathbf{X}_i, Y_i\}$ be jointly stationary processes and $\varphi(\cdot)$ be a Borel measurable function on the real line. This paper aims to give the result of the asymptotic normality of the regression function estimator of

$$m(\mathbf{x},\varphi) = \mathbb{E}[\varphi(Y_1) \mid \mathbf{X}_1 = \mathbf{x}],$$

for $\mathbf{x} \in \mathbb{R}^d$, whenever it exists, i.e.,

$$\mathbb{E}\left(|\varphi(Y_1)|\right) < \infty.$$

Our model includes some special cases of regression models (1.1) depending on the choice the function $\varphi(\cdot)$, in particular the following ones

- $\varphi(Y) = \mathbb{1}\{Y \leqslant y\}$ gives the conditional distribution of Y_1 given $\mathbf{X}_1 = \mathbf{x}$.
- $\varphi(Y) = Y^k$ gives the conditional moments of Y_1 given $\mathbf{X}_1 = \mathbf{x}$.

For more details and motivation one can see the work of [5] and [11]. In the following, we suppose that

$$\mathbb{E}[|\varphi(Y_1)|^p] < \infty, \quad p \geqslant 3, \tag{3.7}$$

and we take in addition, an n-observations $\{\mathbf{X}_i, Y_i\}_{i=1}^n$ random variables from $(\mathbf{X}, Y)$ that is supposed to be stationary and weakly dependent. Similar to the setup in previous works, our estimator of $m(\varphi,\mathbf{x})$ will be obtained by taking the ratio of wavelet estimators of $g(\varphi,\mathbf{x}) = m(\mathbf{x},\varphi)f_{\mathbf{X}}(\mathbf{x})$ and $f_{\mathbf{X}}(\mathbf{x})$, with $f_{\mathbf{X}}(\mathbf{x})$ is an unknown density function of $\mathbf{X}$. Firstly, from (2.2), the linear wavelet estimator of $f_{\mathbf{X}}(\mathbf{x}) \in \mathbf{L}^2(\mathbb{R}^d)$ is introduced by

$$\widehat{f}_{\mathbf{X}}(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{a}_{\tau,\mathbf{k}}\phi_{\tau,\mathbf{k}}(\mathbf{x}) \tag{3.8}$$

or, equivalently, as

$$\widehat{f}_{\mathbf{X}}(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{a}_{j_0\mathbf{k}}\phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{j=j_0}^{\tau}\sum_{i=1}^{N}\sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{b}_{ij\mathbf{k}}\psi_{i,j,\mathbf{k}}(\mathbf{x}), \tag{3.9}$$

with $\widehat{a}_{j_0\mathbf{k}}$ and $\widehat{b}_{ij\mathbf{k}}$ are the unbiased empirical estimates of the coefficients $\{a_{\tau,\mathbf{k}}\}$ and $\{b_{ij\mathbf{k}}\}$ respectively, that is,

$$\widehat{a}_{\tau,\mathbf{k}} = \frac{1}{n}\sum_{i=1}^{n}\phi_{\tau,\mathbf{k}}(\mathbf{X}_i), \text{ and } \widehat{b}_{ij\mathbf{k}} = \frac{1}{n}\sum_{i=1}^{n}\psi_{i,j,\mathbf{k}}(\mathbf{X}_i) \tag{3.10}$$

and for any fixed $j_0 \leqslant \tau$, where $\tau = \tau(n)$ expresses the resolution level as a strictly positive integer depending only on n, tends to infinity at a rate specified below. Remark that the regularity and the compact support conditions on $\phi(\cdot)$ and $\psi(\cdot)$ ensure that the previous summations are finite for any fixed $\mathbf{x}$, which is important in practice. Note that in this case the support of $\phi(\cdot)$ and $\psi_i(\cdot)$ is a monotonically increasing function of their degree of differentiability [17]. Now, notice that $g(\varphi,\cdot) \in \mathbf{L}^2(\mathbb{R}^d)$. It follows that $g(\varphi,\cdot)$ can be represented on a wavelet basis and its linear estimate can be obtained by

$$\widehat{g}_n(\varphi,\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{a}'_{\tau,\mathbf{k}}\phi_{\tau,\mathbf{k}}(\mathbf{x}), \tag{3.11}$$

or, equivalently, as

$$\widehat{g}_n(\varphi,\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{a}'_{j_0\mathbf{k}}\phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{j=j_0}^{\tau}\sum_{i=1}^{N}\sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{b}'_{ij\mathbf{k}}\psi_{i,j,\mathbf{k}}(\mathbf{x}). \tag{3.12}$$

The coefficients estimators are given respectively by

$$\widehat{a}'_{\tau,\mathbf{k}} = \frac{1}{n}\sum_{i=1}^{n}\varphi(Y_i)\phi_{\tau,\mathbf{k}}(\mathbf{X}_i), \tag{3.13}$$

and

$$\widehat{b}'_{ij\mathbf{k}} = \frac{1}{n}\sum_{i=1}^{n}\varphi(Y_i)\psi_{i,j,\mathbf{k}}(\mathbf{X}_i), \tag{3.14}$$

for any $j_0 \leqslant \tau$. From the practical point of view, the multiresolution properties of the wavelets provide an implementation of an estimation algorithm convenient in computation, with an important saving of memory. Indeed, the bandwidth for this strategy is essential, minimized to the form 2^j for j easy to be selected and only a small number of values of j (say three or four) need to be considered in practice, [34].

4. Assumptions and main results

In this section, we state our results on the asymptotic normality of the wavelets regression estimator with weak-dependent data. Below, we introduce general forms of the Nadaraya-Watson kernel estimator of $m(\cdot,\varphi)$ ([42] and [55]), and of the Akaike-Parzen-Rosenblatt kernel density estimator of $f_{\mathbf{X}}(\cdot)$ ([1], [49] and [44]). Let the kernel $K(\cdot)$ be any function satisfying some regularity conditions and $(h_n)_{n\geqslant 1}$ be a sequence of positive constants converging to zero and

$$nh_n^d \longrightarrow \infty \text{ as } n \longrightarrow \infty.$$

The kernel-type estimator of the density function $f_{\mathbf{X}}(\cdot)$ of $\mathbf{X}$ is given, for $\mathbf{x} \in \mathbb{R}^d$, by

$$f_{n,h_n}(\mathbf{x}) := \frac{1}{nh_n^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right). \tag{4.1}$$

[44] has shown, under some assumptions on $K(\cdot)$, that $f_{n,h_n}(\cdot)$ is an asymptotically unbiased and consistent estimator for $f_{\mathbf{X}}(\cdot)$ whenever $h_n \rightarrow 0$, $nh_n^d \rightarrow \infty$ and $\mathbf{x}$ is a continuity point of $f_{\mathbf{X}}(\cdot)$. Under some additional assumptions on $f_{\mathbf{X}}(\cdot)$ and h_n, he obtained an asymptotic normality result, too. The general kernel-type estimator of $m(\cdot,\varphi)$ is given, for $\mathbf{x} \in \mathbb{R}^d$, by

$$\widehat{m}_{n;h_n}(\mathbf{x},\varphi) := \frac{\sum_{i=1}^{n} \varphi(Y_i) K((\mathbf{x}-\mathbf{X}_i)/h_n)}{\sum_{i=1}^{n} K((\mathbf{x}-\mathbf{X}_i)/h_n)}. \tag{4.2}$$

By setting $\varphi(y) = y$ (or $\varphi(y) = y^k$) into (4.2), we get the classical Nadaraya-Watson kernel regression function estimator of $m(\mathbf{x}) := \mathbb{E}(Y \mid \mathbf{X}=\mathbf{x})$ given by

$$\widehat{m}_{n;h_n}(\mathbf{x}) := \frac{\sum_{i=1}^{n} Y_i K((\mathbf{x}-\mathbf{X}_i)/h_n)}{\sum_{i=1}^{n} K((\mathbf{x}-\mathbf{X}_i)/h_n)}, \tag{4.3}$$

or

$$\widehat{m}_{n;h_n}(\mathbf{x}) := \frac{\sum_{i=1}^{n} Y_i^k K((\mathbf{x}-\mathbf{X}_i)/h_n)}{\sum_{i=1}^{n} K((\mathbf{x}-\mathbf{X}_i)/h_n)}. \tag{4.4}$$

[42] established similar results to those of [44] for $\widehat{m}_{n;h_n}(\mathbf{x})$ as an estimator for $\mathbb{E}(Y \mid \mathbf{X}=\mathbf{x})$. By setting $\varphi_t(y) = \mathbb{1}[y \leqslant t]$, for $t \in \mathbb{R}$, into (4.2), we obtain the kernel estimator of the conditional distribution function $F(t|\mathbf{x}) := \mathbb{P}(Y \leqslant t \mid \mathbf{X}=\mathbf{x})$ given by

$$\widehat{F}_{n;h_n}(t|\mathbf{x}) := \frac{\sum_{i=1}^{n} \mathbb{1}(Y_i \leqslant t) K((\mathbf{x}-\mathbf{X}_i)/h_n)}{\sum_{i=1}^{n} K((\mathbf{x}-\mathbf{X}_i)/h_n)}. \tag{4.5}$$

These examples motivate the introduction of the function $\varphi(\cdot)$ in our setting, we refer to [5] for details. Return to our main concerns with the wavelet type estimators. For $\mathbf{u},\mathbf{v}\in\mathbb{R}^d$, we define the kernel $K(\mathbf{u},\mathbf{v})$ by

$$K(\mathbf{u},\mathbf{v}) := \sum_{\mathbf{k}\in\mathbb{Z}^d}\phi(\mathbf{u}-\mathbf{k})\phi(\mathbf{v}-\mathbf{k}). \tag{4.6}$$

Using the fact that

$$|\phi(\mathbf{x})| \leqslant \frac{A_{d+1}}{(1+\|\mathbf{x}\|)^{d+1}},$$

we infer that the kernel function $K(\cdot,\cdot)$ defined in (4.6) converges uniformly in $\mathbf{u},\mathbf{v}\in\mathbb{R}^d$, in the sense that

$$\sup_{\mathbf{u},\mathbf{v}\in\mathbb{R}^d}\left|\sum_{\mathbf{k}\in\mathbb{Z}^d}\phi(\mathbf{u}-\mathbf{k})\phi(\mathbf{v}-\mathbf{k})\right| < \infty,$$

and satisfies, for any $j\geqslant 1$, ([41], page 33)

$$|K(\mathbf{v},\mathbf{u})| \leqslant \frac{C_j}{(1+\|\mathbf{v}-\mathbf{u}\|)^j}, \tag{4.7}$$

for some constant C_j. From (4.8), it follows that

$$\int_{\mathbb{R}^d}|K(\mathbf{v},\mathbf{u})|^j d\mathbf{v} \leqslant G_j(d), \tag{4.8}$$

where

$$G_j(d) = 2\pi^{d/2}\frac{\Gamma(d)\Gamma(j+d(j-1))}{\Gamma(d/2)\Gamma((d+1)j)}C_{d+1}^j$$

and $\Gamma(t)$ is the Gamma function, that is,

$$\Gamma(t) := \int_0^\infty y^{t-1}\exp(-y)\,dy.$$

Furthermore, we have with assumption given on $\phi(\cdot)$, for $|\beta|=1$, ([41], p. 33)

$$\left|\frac{\partial K(\mathbf{u},\mathbf{v})}{\partial \mathbf{u}_i}\right| \leqslant \frac{C_2}{(1+\|\mathbf{u}-\mathbf{v}\|)^2}, \quad i=1,\ldots,d. \tag{4.9}$$

By combining (3.8), (3.10) and (4.6), we observe that the linear estimates of the regression function $\widehat{m}_n(\varphi,\mathbf{x})$ can be written as an extended kernel estimator for

$$K_{h_n,\mathbf{x}}(\mathbf{X}_i) = \frac{1}{h_n^d}K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{X}_i}{h_n}\right) \text{ and } h_n = 2^{-\tau(n)},$$

as follows

$$\widehat{m}_n(\varphi,\mathbf{x}) = \begin{cases} \dfrac{\widehat{g}_n(\varphi,\mathbf{x})}{\widehat{f}_{\mathbf{X}}(\mathbf{x})} = \dfrac{\widehat{m}_{2,n}(\varphi,\mathbf{x})}{\widehat{m}_{1,n}(\mathbf{x})}, \text{ if } \widehat{f}_n(\mathbf{x})\neq 0, \\[2ex] \dfrac{1}{n}\displaystyle\sum_{i=1}^n\varphi(Y_i), \quad \text{otherwise.}\end{cases}$$

for

$$\widehat{m}_{1,n}(\varphi,\mathbf{x}) = \frac{\sum_{i=1}^{n} K_{h_n,\mathbf{x}}(\mathbf{X}_i)}{n\mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_1)}.$$

$$\widehat{m}_{2,n}(\varphi,\mathbf{x}) = \frac{\sum_{i=1}^{n} \varphi(Y_i)K_{h_n,\mathbf{x}}(\mathbf{X}_i)}{n\mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_1)}.$$

Our main result concerning the asymptotic normality of the wavelet regression estimator is stated under the following assumptions:

(A.1) Let $\{\mathbf{X}_i, Y_i\}_{i\geqslant 1}$ be an $\mathbb{R}^d \times \mathbb{R}$-valued stationary and $\widetilde{\alpha}$-weakly dependent process defined on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ such that the dependence coefficient satisfies

$$\sum_{i=1}^{\infty} i^{1+\frac{1}{d}} \tilde{\alpha}^{\frac{1}{2}}(i) < \infty.$$

(A.2) (i) The density function $f_{\mathbf{X}}(\cdot)$ of $\mathbf{X}_i$ is continuous and, for all $\mathbf{x} \in \mathbb{R}^d$, there exists M a strictly positive constant such that

$$0 < f_{\mathbf{X}}(\mathbf{x}) \leqslant M.$$

(ii) Let $f_{\mathbf{X}_i,\mathbf{X}_j}(\cdot)$ be the joint density function of the pair $(\mathbf{X}_i, \mathbf{X}_j)$ satisfying

$$\sup_{(\mathbf{x},\mathbf{y})\in[-2L,2L]^{2d}} f_{\mathbf{X}_i,\mathbf{X}_j}(\mathbf{x},\mathbf{y}) = M < \infty.$$

(A.3) The conditional variance $\sigma_\varphi^2(\mathbf{x}) = \mathbb{E}((\varphi(Y_1) - m(\varphi,\mathbf{x}))^2 | \mathbf{X}_1)$ exists and continuous in some neighborhood of $\mathbf{x} \in \mathbb{R}^d$ in the sense that

$$\sup_{\{\mathbf{u}:\|\mathbf{x}-\mathbf{u}\|\leqslant h\}} \left|\sigma_\varphi^2(\mathbf{u}) - \sigma_\varphi^2(\mathbf{x})\right| = o(1) \quad \text{as} \quad h \to 0.$$

(A.4) The regression function $m(\cdot,\varphi)$ is continuous in some neighborhood of $\mathbf{x} \in \mathbb{R}^d$, that is

$$\sup_{\{\mathbf{u}:\|\mathbf{x}-\mathbf{u}\|\leqslant h\}} |m(\varphi,\mathbf{u}) - m(\varphi,\mathbf{x})| = o(1) \quad \text{as} \quad h \to 0.$$

(A.5) Let $p = p(n)$, $q = q(n)$ be an integer valued sequences tending to ∞ with n such q is greater than p and satisfy $p^2 + q^2 = o(n)$. Put $k := k(n) = \left[\frac{n}{p+q}\right] \to \infty$, such that $\frac{kq}{n} \to 1$ and

$$\frac{k}{\sqrt{n}} \tilde{\alpha}^{\frac{1}{2}}(p) \to 0. \tag{4.10}$$

(A.6) (i) The multiresolution analysis is r-regular.

(ii) The density $f \in \mathbf{B}_{s,p,q}$ for some $0 < s < r, 1 \leqslant p,q \leqslant \infty$.

(iii) The function $g \in \mathbf{B}_{s,p,q}$ for some $0 < s < r, 1 \leqslant p,q \leqslant \infty$.

Comments on the assumptions: The assumption **(A.1)** involves the stationary and the relaxed condition of $\widetilde{\alpha}$-weakly dependent for the process $(\mathbf{X}_i, Y_i)$. The condition on the coefficient $\tilde{\alpha}$ is equivalent to the standard assumption of the Riemaniann decay $\tilde{\alpha}(r) = O(r^{-\lambda}), \lambda > 4 + \frac{2}{d}$. Note that for an exponential decay assumption on the coefficient $\tilde{\alpha}(r) = O(\exp(-\lambda r), \lambda > 0$, the condition is automatically satisfied. The assumption **(A.2)** are mild conditions on the joint and marginal probability distribution and the joint density function of the process $(\mathbf{X}_i, Y_i)_{i\in\mathbb{Z}}$, commonly used in the nonparametric curves estimation literature. We mention that the smoothness conditions **(A.3)** and **(A.4)** are similar to those in [5] and [12]. These later assumptions play an important role in studying the asymptotic variance terms. The assumption **(A.5)** is linked to Bernstein's blocking approach that we need in the proof. If we take $q = n^{\gamma}$ and $p = n^{\beta}$, for $0 < \beta < \gamma$ and $\gamma \in]0, \frac{1}{2}[$, clearly we have $p^2 + q^2 = o(n)$ and $k = O(n^{1-\gamma}) \to \infty$. So, $\frac{kq}{n} \to 1$. One can see that when $\tilde{\alpha}$ has the geometric decay, the condition (4.10) is fulfilled without any additional requirements. On the other hand, for the Riemanian case, i.e., $(\tilde{\alpha}(j) = O(j^{-\lambda}),\ \lambda > 4 + \frac{2}{d})$, if we take $2\gamma + \beta\lambda > 1$ the condition (4.10) is equivalent to

$$\frac{k}{\sqrt{n}}\tilde{\alpha}^{\frac{1}{2}}(p) = O\left(\frac{1}{n^{-\frac{1}{2}+\gamma+\frac{\beta\lambda}{2}}}\right) = o(1).$$

Hence the Theorem 4.1 conditions can be verified for a reasonably choice of the parameters γ and β. We may take as example, $\beta = \frac{1}{4}, \gamma = \frac{1}{3}$. The assumption **(A.6)** is important to approximate the bias on the generalized space of Besov ($\mathbf{B}_{s,p,q}$ for $s > d/p$).

Below, we write $Z \xrightarrow{\mathscr{D}} \mathscr{N}(\mu, \sigma^2)$ whenever the random variable Z follows a normal law with expectation μ and variance σ^2. Our main result is summarized in the following theorem.

THEOREM 4.1. *Assume that*

$$\tau = \tau_n \to \infty, \quad \sqrt{n 2^{-(d+2\beta_1)\tau}} \to 0 \ \text{as} \ n \to \infty.$$

*If the Assumptions **(A.1)–(A.6)** are verified, then we have*

$$\sqrt{n2^{-d\tau}}(\widehat{m}_n(\varphi, \mathbf{x}) - m(\varphi, \mathbf{x})) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \Sigma_{\varphi}^2),$$

where $\xrightarrow{\mathscr{D}}$ *means the convergence in distribution and*

$$\Sigma_{\varphi}^2 := \Sigma_{\varphi}^2(\mathbf{x}) := \frac{\sigma_{\varphi}^2(\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})} \int_{\mathbb{R}^d} \left(\sum_{\mathbf{k}\in\mathbb{Z}^d} \phi(\mathbf{0}+\mathbf{k})\phi(\mathbf{v}+\mathbf{k}) \right)^2 d\mathbf{v},$$

where $\sigma_{\varphi}^2(\mathbf{x})$ *is given in **(A.3)**.*

An application of the Slutsky theorem gives the following corollary.

COROLLARY 4.1. *Assume that*

$$\tau = \tau_n \to \infty, \quad \sqrt{n2^{-(d+2\beta_1)\tau}} \to 0 \ \text{as} \ n \to \infty.$$

*If the Assumptions **(A.1)–(A.6)** are verified, then we have*

$$\sqrt{n2^{-d\tau}}\widehat{\Sigma}_{\varphi}^{-1}(\widehat{m}_n(\varphi,\mathbf{x}) - m(\varphi,\mathbf{x})) \xrightarrow{\mathscr{D}} \mathscr{N}(0,1),$$

where $\widehat{\Sigma}_{\varphi}^{2}$ *is given in (4.12).*

From Theorem 4.1 and using the same reasoning of [5], we have the following corollary.

COROLLARY 4.2. *Assume that*

$$\tau = \tau_n \to \infty, \quad \sqrt{n2^{-(d+2\beta_1)\tau}} \to 0 \ \text{as} \ n \to \infty.$$

*If the Assumptions **(A.1)–(A.6)** are verified, then we have*

$$\sqrt{n2^{-d\tau}}(\widehat{f}_{\mathbf{X}}(\mathbf{x}) - f_{\mathbf{X}}(\mathbf{x})) \xrightarrow{\mathscr{D}} \mathscr{N}\left(0,\Sigma^{*2}\right),$$

where

$$\Sigma^{*2} := \Sigma^{*2}(\mathbf{x}) := f_{\mathbf{X}}(\mathbf{x}) \int_{\mathbb{R}^d} \left(\sum_{\mathbf{k}\in\mathbb{Z}^d} \phi(\mathbf{0}+\mathbf{k})\phi(\mathbf{v}+\mathbf{k}) \right)^2 d\mathbf{v}.$$

REMARK 4.1. The conclusion of Theorem 4.1 is similar to the result of [27], for the uni-dimensional Parzen-Rosenblatt kernel density estimator defined in the Hölder space, under the notion of Ψ-weak dependence of $\psi := c\mathrm{Lip} f \mathrm{Lip} g$, (resp. $\Psi := \min(\mathrm{Lip} f, \mathrm{Lip} g)$) with an arithmetic decay of the dependence coefficients $\theta(r) = O(r^{-(12+\delta)})$, (resp. $\theta(r) = O(r^{-(9+\delta)})$), for some strictly positive constant δ.

REMARK 4.2. For some constants $\delta > 0$ and $p \geqslant 2$, if $\mathbb{E}|X_j|^p < \infty$, we have, by Markov inequality

$$\mathbb{P}(|X| > \varepsilon) \leqslant \frac{\mathbb{E}|X|^p}{\varepsilon^p} < \left(\frac{C}{\varepsilon}\right)^{p+\delta},$$

implies that $Q_X(u) \leqslant Cu^{\frac{-1}{p+\delta}}$, and for any positive constant δ and $t \in [0,1]$, we have

$$\int_0^t Q_X^p(u) < C\int_0^t u^{-\frac{p}{p+\delta}}du = \frac{p+\delta}{\delta}Ct^{1-\frac{p}{p+\delta}} = C'_{p,\delta}t^{1-\frac{p}{p+\delta}}. \tag{4.11}$$

4.1. Confidence interval

The asymptotic variance Σ_{φ}^2 in the central limit theorem depends on the unknown functions (of the conditional variance of Y given $\mathbf{X}$ and the density function $f_{\mathbf{X}}(\cdot)$ of $\mathbf{X}$) should be estimated in practice. The estimation requires a choice of some bounded compactly supported family of discrete wavelets founded from literature (such as the commonly used [17] wavelets) for multiresolution level τ_n large enough and an adaptive initial level j_0. We select the unknown parameters through the wavelet methods and plug-in approach. Using the estimators (3.8) we establish $\widehat{\Sigma}_{\varphi}^2$ a consistent estimate of the variance Σ_{φ}^2 given by

$$\widehat{\Sigma}_{\varphi}^2 := \frac{\widehat{\sigma}_{\varphi}^2(\mathbf{x})}{\widehat{f}_{\mathbf{X}}(\mathbf{x})} \int_{\mathbb{R}^d} \left(\sum_{\mathbf{k}\in\mathbb{Z}^d} \phi(\mathbf{0}+\mathbf{k})\phi(\mathbf{v}+\mathbf{k}) \right)^2 d\mathbf{v}, \tag{4.12}$$

where $\widehat{\sigma}_{\varphi}^2(\mathbf{x})$ is the plug-in estimate of the condition variance σ_{φ}^2 from the observed data, that is

$$\widehat{\sigma}_{\varphi}^2(\mathbf{x}) = \frac{\widehat{m}_{3,n}(\varphi,\mathbf{x})}{\widehat{m}_{1,n}(\mathbf{x})},$$

where

$$\widehat{m}_{3,n}(\varphi,\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{a}''_{j_0\mathbf{k}}\phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{j=j_0}^{\tau}\sum_{i=1}^{N}\sum_{\mathbf{k}\in\mathbb{Z}^d} \widehat{b}''_{ij\mathbf{k}}\psi_{i,j,\mathbf{k}}(\mathbf{x}). \tag{4.13}$$

The considered coefficients estimators are given respectively, for any $j_0 \leqslant \tau$, by

$$\widehat{a}''_{\tau,\mathbf{k}} = \frac{1}{n}\sum_{i=1}^{n}\left\{\varphi(Y_i) - \widehat{m}_n(\varphi,\mathbf{x})\right\}^2 \phi_{\tau,\mathbf{k}}(\mathbf{X}_i),$$

and

$$\widehat{b}''_{ij\mathbf{k}} = \frac{1}{n}\sum_{i=1}^{n}\left\{\varphi(Y_i) - \widehat{m}_n(\varphi,\mathbf{x})\right\}^2 \psi_{i,j,\mathbf{k}}(\mathbf{X}_i).$$

Furthermore, from (3.11) and (4), the approximate confidence intervals of $m(\varphi,\mathbf{x})$ can be obtained as

$$m(\varphi,\mathbf{x}) \in \left[\widehat{m}_n(\varphi,\mathbf{x}) \pm c_{\alpha}\frac{\widehat{\Sigma}_{\varphi}}{\sqrt{nh_n^d}}\right]$$

where c_{α}, denotes the $(1-\frac{\alpha}{2})$-quantile of the standard normal condition.

REMARK 4.3. In this paper, we have considered the confidence interval application. Since main the result is focused on the pointwise asymptotic normality, the resulting confidence interval is pointwise too. However, construction of confidence bands tends to be challenging, especially for complex nonparametric models (we refer to [54], [33], [31] as general references on confidence bands in nonparametric statistical models). Despite the rich literature on consistent estimation of nonparametric regression, the literature on uniform confidence bands for nonparametric regression by the wavelet methods is limited. [30] have investigated a linear wavelet estimator and proved strong consistency when the level of smoothing, i.e., the level of approximation of $\mathbf{L}^2(\mathbb{R}^d)$ by such multiscales, is allowed to range in some interval depending on n. To obtain a uniform limit theorem with respect to such parameters, the authors make use of modern empirical process theory. As a statistical application, [30] proved that essentially the same limit theorems can be obtained for the hard thresholding wavelet estimator introduced by [24]. In a similar spirit, [56] presented the rates of uniform strong consistency of wavelet estimation for nonparametric function in sup-norm loss by introducing an empirical process approach. [29] provided new uniform rate results for kernel estimators of absolutely regular stationary processes that are uniform in the bandwidth and in infinite-dimensional classes of dependent variables and regressors. It will be of interest to extend the last findings to our framework. The proof of such a statement, however, should require a different methodology than that used in the present paper, and we leave this problem open for future research.

5. Concluding remarks

The present work is mainly concerned with the convergence in distribution to Gaussian processes as well as the construction of confidence intervals of the multivariate wavelets regression estimators from a stationary dependent $\mathbb{R}^d$-process. The dependence assumption on the sequence of random variables is relaxed by using the concept of weak dependence condition given by [26], which takes advantage of covering large classes of interesting models used in econometrics that classical mixing properties can fail to hold as Bernoulli shifts, Markov processes, among many others. It is well known that wavelet estimators outperform kernel ones in representing discontinuities. However, for inhomogeneous curves (i.e., case when $1 \leqslant p < 2$) or that with unknown regularity, the linear smoother methods present some drawbacks to reach the optimal minimax rate of convergence. To circumvent these problems, one can use the nonlinear wavelets of soft or hard thresholding estimators, which requires nontrivial mathematical development and we leave this problem open for forthcoming research. It will be of interest to consider the problem of the conditional U-statistics, investigated by [10], [9], [8], [13], using the wavelet estimation.

6. Mathematical development

This section is devoted to the proofs of our results. The previously presented notation continues to be used in the following.

Proof of Theorem 4.1

Keeping in mind that the setting of Section 2, the kernel $K(\cdot,\cdot)$ is compactly supported satisfiying

$$K(\mathbf{u},\mathbf{v}) = 0 \ \ as \ \ |u_i - v_i| > 2L, \text{ for } i = 1,\ldots,d,$$

and

$$\int_{\mathbb{R}^d} K(\mathbf{x},\mathbf{x}+\mathbf{u}) = \int_{[-L,L]^d} K(\mathbf{0},\mathbf{u})d\mathbf{u} = 1. \tag{6.1}$$

We have, by assumption **(A.2)**(i) and equation (6.1), for n large enough, we have

$$\begin{aligned}\mathbb{E}\widehat{f}_n(\mathbf{x}) = \mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_1) &= \int_{[-2L,2L]^d} K\left(\frac{\mathbf{x}}{h_n}, \frac{\mathbf{x}}{h_n}+\mathbf{u}\right) f(\mathbf{x}+\mathbf{u}h_n)d\mathbf{u}\\ &= f(\mathbf{x})+o(1) > 0. \end{aligned}\tag{6.2}$$

Define

$$\widehat{m}_{1,n}(\mathbf{x}) = \frac{\sum_{i=1}^{n} K_{h_n,\mathbf{x}}(\mathbf{X}_i)}{n\mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_1)} = \frac{\widehat{f}_{\mathbf{X},n}(\mathbf{x})}{\mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_i)},$$

$$\widehat{m}_{2,n}(\varphi,\mathbf{x}) = \frac{\sum_{i=1}^{n} \varphi(Y_i)K_{h_n,\mathbf{x}}(\mathbf{X}_i)}{n\mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_1)} = \frac{\widehat{g}_n(\varphi,\mathbf{x})}{\mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_i)}.$$

We consider the following decomposition

$$\begin{aligned}\widehat{m}_n(\varphi,\mathbf{x}) - m(\varphi,\mathbf{x}) &= \widehat{m}_n(\varphi,\mathbf{x}) - \mathbb{E}\widehat{m}_n(\varphi,\mathbf{x}) + \mathbb{E}\widehat{m}_n(\varphi,\mathbf{x}) - m(\varphi,\mathbf{x})\\ &= \frac{Q_n(\mathbf{x})+U_n(\mathbf{x})}{\widehat{m}_{1,n}(\mathbf{x})} + B_n(\mathbf{x}),\end{aligned}$$

where

$$\begin{aligned}Q_n(\mathbf{x}) &= \widehat{m}_{2,n}(\mathbf{x}) - \mathbb{E}\widehat{m}_{2,n}(\mathbf{x}) - m(\varphi,\mathbf{x})(\widehat{m}_{1,n}(\mathbf{x}) - \mathbb{E}\widehat{m}_{1,n}(\mathbf{x})),\\ U_n(\mathbf{x}) &= B_n(\mathbf{x})(\mathbb{E}\widehat{m}_{1,n}(\mathbf{x}) - \widehat{m}_{1,n}(\mathbf{x})),\\ B_n(\mathbf{x}) &= \frac{(\mathbb{E}\widehat{g}_n(\mathbf{x}) - g(\mathbf{x})) - m(\mathbf{x})\left(\mathbb{E}\widehat{f}_n(\mathbf{x}) - f(\mathbf{x})\right)}{\mathbb{E}\widehat{f}_n(\mathbf{x})}.\end{aligned}$$

In what follows, we obtain our main result of Theorem 4.1 from some lemmas establishing respectively, the asymptotic convergence in probability (noted $o_{\mathbb{P}}(1)$) of the term $\widehat{m}_{1,n}(\mathbf{x})$, asymptotic convergence in probability of the bias term and the asymptotic normality of $Q_n(\mathbf{x})$. Using the relation (6.2), we have the following lemma.

LEMMA 6.1. *Under the assumption **(A.6)** and*

$$nh_n^{d+2(s-\frac{d}{p})} \to 0, \;\; as \;\; n\to\infty, \;\; s > d/p,$$

we have

$$\sqrt{nh_n^d}B_n(\mathbf{x}) = O\left(\sqrt{nh_n^{d+2(s-\frac{d}{p})}}\right) = o(1).$$

Proof of Lemma 6.1. The study of $B_n(x)$ of the regression function $m(\varphi,\mathbf{x})$ is purely analytical and does not depend on the dependence properties of the sequences $(\mathbf{X}_i)_{i\in\mathbb{Z}}$. Following [39, Lemma.2], the assumptions given in **(A.6)** satisfy, for $s>d/p$, $h_n := 2^{-j(n)}$,

$$\sup_{\mathbf{x}\in\mathbb{R}^d}\left|\mathbb{E}\widehat{f}_n(\mathbf{x})-f(\mathbf{x})\right| \leqslant Ch_n^{(s-d/p)}J_{s,p,q}(f),$$

$$\sup_{\mathbf{x}\in\mathbb{R}^d}\left|\mathbb{E}\widehat{g}_n(\mathbf{x})-g(\mathbf{x})\right| \leqslant C_1h_n^{(s-d/p)}J_{s,p,q}(g),$$

and $\sup_{\mathbf{x}\in\mathbb{R}^d}|g(\mathbf{x})|<\infty$. Thus from (6.2) we deduce the following bound

$$\begin{aligned}|B_n(\mathbf{x})| &\leqslant \frac{1}{|\mathbb{E}K_{h_n,\mathbf{x}}(\mathbf{X}_1)|}\left[\sup_{\mathbf{x}\in\mathbb{R}^d}|\mathbb{E}\widehat{g}_n(\mathbf{x})-g(\mathbf{x})|+\left|\frac{g(\mathbf{x})}{f_{\mathbf{X}_i(\mathbf{x})}}\right|\sup_{\mathbf{x}\in\mathbb{R}^d}\left|\mathbb{E}\widehat{f}_n(\mathbf{x})-f(\mathbf{x})\right|\right]\\ &= O\left(h_n^{(s-d/p)}\right). \end{aligned}\tag{6.3}$$

The upper bound result in (6.3) with assumptions of Lemma 6.1, allow to obtain

$$\sqrt{nh_n^d}B_n(\mathbf{x}) = o(1).$$

Hence the proof is completed. $\square$

REMARK 6.1. Note that for a choice of $h_n := 2^{-\tau_n} \asymp n^{-\theta}$ for some $\frac{1}{d+2(s-\frac{d}{p})} < \theta \leqslant \frac{1}{d}$, we conclude that

$$h_n \to 0,\ \sqrt{nh_n^{d+2(s-\frac{d}{p})}} \to 0 \ \text{ and } \ nh_n^d \to \infty$$

.

LEMMA 6.2. *Under assumptions **(A.1)** and **(A.2)**, we have*

$$\widehat{m}_{1,n}(\mathbf{x})-1 = o_{\mathbb{P}}(1).$$

Proof of Lemma 6.2. We can see that $\mathbb{E}(\widehat{m}_{1,n}(\mathbf{x}))=1$, then by Chebyshev's inequality, it is sufficient to show that $\operatorname{Var}(\widehat{m}_{1,n}(\mathbf{x}))\to 0$. Note that

$$\begin{aligned}\operatorname{Var}(\widehat{m}_{1,n}(\mathbf{x})) &= \operatorname{var}\left(\frac{1}{n\mathbb{E}(K_{h_n,\mathbf{x}}(\mathbf{X}_i))}\sum_{i=1}^n K_{h_n,\mathbf{x}}(\mathbf{X}_i)\right)\\ &= \frac{1}{n\mathbb{E}^2(K_{h_n,\mathbf{x}}(\mathbf{X}_1))}\operatorname{Var}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_1)\right)\\ &\quad+\frac{2}{n^2\mathbb{E}^2(K_{h_n,\mathbf{x}}(\mathbf{X}_1))}\sum_{\substack{i,j=1\\ i<j}}^n \operatorname{Cov}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_i),K_{h_n,\mathbf{x}}(\mathbf{X}_j)\right)\\ &= Q_1+Q_2. \end{aligned}\tag{6.4}$$

The right term Q_1 may be sub bounded as following

$$\begin{aligned}\frac{1}{n\mathbb{E}^2(K_{h_n,\mathbf{x}}(\mathbf{X}_1))}\mathrm{Var}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_1)\right) &= \frac{C}{n}\mathbb{E}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_1)\right)^2 - \frac{1}{n}\\ &= \frac{C}{n}\int_{\mathbb{R}^d} K^2\left(\frac{\mathbf{u}}{h_n},\mathbf{u}\right) f_{\mathbf{X}}(h_n\mathbf{u})\,d\mathbf{u} + o(1)\\ &< \frac{C}{n}MG_2(d) = o(1). \end{aligned} \tag{6.5}$$

Next, we evaluate the term Q_2 on the right side of (6.4). The stationary assumption on $\mathbf{X}$ in combination with (6.2), we readily infer that

$$\begin{aligned} Q_2 &= \frac{2}{n^2\mathbb{E}^2(K_{h_n,\mathbf{x}}(\mathbf{X}_i))}\sum_{j=2}^{n}(n-j)\mathrm{Cov}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_1), K_{h_n,\mathbf{x}}(\mathbf{X}_j)\right)\\ &\leqslant \frac{2}{n\mathbb{E}^2(K_{h_n,\mathbf{x}}(\mathbf{X}_i))}\sum_{j=2}^{n}\left|\mathrm{Cov}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_1), K_{h_n,\mathbf{x}}(\mathbf{X}_j)\right)\right|\\ &= \frac{C}{n}\left[\sum_{j=2}^{c_n} + \sum_{j=c_n+1}^{n}\right]\left|\mathrm{Cov}(K_{h_n,\mathbf{x}}(\mathbf{X}_1), K_{h_n,\mathbf{x}}(\mathbf{X}_j))\right|\\ &:= Q_{2,1} + Q_{2,2}, \end{aligned} \tag{6.6}$$

where c_n be an integer sequence such that $c_n = o(n)$ tends to infinity. Notice that we have

$$\begin{aligned} &\mathrm{Cov}(K_{h_n,\mathbf{x}}(\mathbf{X}_1), K_{h_n,\mathbf{x}}(\mathbf{X}_j))\\ &= \frac{1}{h_n^{2d}}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{u}}{h_n}\right)K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{v}}{h_n}\right) f_{\mathbf{X_1},\mathbf{X_2}}(\mathbf{u},\mathbf{v})d\mathbf{u}d\mathbf{v}\\ &\quad - \frac{1}{h_n^{2d}}\left(\int_{\mathbb{R}^d} K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{u}}{h_n}\right) f_{\mathbf{X}}(\mathbf{u})d\mathbf{u}\right)^2\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{u}\right)K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{v}\right) f_{\mathbf{X_1},\mathbf{X_2}}(\mathbf{x}+h_n\mathbf{u},\mathbf{x}+h_n\mathbf{v})d\mathbf{u}d\mathbf{v} + o(1). \end{aligned}$$

An application of Lebesgue dominated convergence theorem gives

$$\begin{aligned} &\left|\int_{[-2L,2L]^d}\int_{[-2L,2L]^d} K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{u}\right)K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{v}\right) f_{\mathbf{X_1},\mathbf{X_2}}(\mathbf{x}+h_n\mathbf{u},\mathbf{x}+h_n\mathbf{v})d\mathbf{u}d\mathbf{v}\right|\\ &\qquad = f_{\mathbf{X_1},\mathbf{X_2}}(\mathbf{x},\mathbf{x})\left|\left(\int_{[-2L,2L]^d} K\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{u}\right)d\mathbf{u}\right)^2\right| \leqslant M'G_1^2(d). \end{aligned}$$

Thus, we infer that

$$\left|\mathrm{Cov}(K_{h_n,\mathbf{x}}(\mathbf{X}_1), K_{h_n,\mathbf{x}}(\mathbf{X}_j))\right| \leqslant M'G_1^2(d),$$

and

$$Q_{2,1} = O\left(\frac{c_n}{n}\right) = o(1). \tag{6.7}$$

The second term $Q_{2.2}$ on the right hand of (6.6) is bounded using the weak dependence condition. Recall that, for any $j \in \mathbb{Z}$, $\mathbf{X}_j = (X_{j_1}, \ldots, X_{jd})$ is an $\mathbb{R}^d$-valued process satisfying the $\tilde{\alpha}$-dependence. By stationary and the fact that kernel $K_{h_n,\mathbf{x}}(\mathbf{X}_i)$ is bounded and Lipschitz continuous function in $\mathbb{R}^d \to \mathbb{R}$, with

$$\mathrm{Lip} K_{h_n,\mathbf{x}}(\mathbf{X}_i) = O\left(\frac{1}{h_n^{d+1}}\right),$$

also we have by (4.7),

$$\|K_{h_n,\mathbf{x}}\|_\infty \leqslant \frac{1}{h_n^d} C_{d+1}.$$

Using the relation (3.2), we get the following bound

$$\begin{aligned} Q_{2,2} &= \frac{C}{n} \sum_{j=c_n+1}^{n} \left|\mathrm{Cov}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_1), K_{h_n,\mathbf{x}}(\mathbf{X}_j)\right)\right| \\ &\leqslant \frac{C}{n h_n^{2d+1}} \sum_{i=c_n}^{n} \tilde{\alpha}^{\frac{1}{2}}(i) < \frac{C}{c_n^{\nu+1} h_n^{2d+1}} \sum_{i=c_n}^{n} i^\nu \tilde{\alpha}^{\frac{1}{2}}(i). \end{aligned}$$

If we select $c_n = h_n^{-d} = o(n)$, we have $c_n^{\nu+1} h_n^{2d+1} = 1$ provided that $\nu = 1 + \frac{1}{d}$. By assumption (**A.1**), we get

$$Q_{2,2} = o(1).$$

This when combined with the results (6.5) and (6.7), lead to

$$\mathrm{Var}(\widehat{m}_{1,n}) \to 0.$$

This permits to conclude that $\widehat{m}_{n,1}(\mathbf{x})$ converges in quadratic mean to $\mathbb{E}(\widehat{m}_{1,n}) = 1$. This implies the probability convergence expressed in the Lemma 6.2 as $n \to \infty$. □

The above results given in Lemma 6.2 and Lemma 6.1 mean that $B_n(\mathbf{x})$ and $U_n(\mathbf{x})$ are asymptotically negligible, as $n \to \infty$. Hence, to prove the result of the Theorem 4.1, we should only consider the Lemma 6.3 for $Q_n(\mathbf{x})$.

LEMMA 6.3. *If the assumptions **(A.1)–(A.5)** are verified, we have, as $n \to \infty$*

$$\sqrt{n h_n^d} Q_n(\mathbf{x}) \xrightarrow{\mathscr{D}} \mathscr{N}\left(0, \Sigma_\varphi^2\right),$$

where Σ_φ^2 is defined in Theorem 4.1.

Proof of Lemma 6.3. Let us introduce the following notation

$$\xi_{h_n i} := \sqrt{h_n^d}(\varphi(Y_i) - m(\varphi, \mathbf{x})) K_{h_n,\mathbf{x}}(\mathbf{X}_i),$$

and

$$\sqrt{n h_n^d} Q_n(\mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{n,i},$$

where

$$Z_{n,i} = \frac{1}{\mathbb{E}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_i)\right)}\left(\xi_{h_n,i} - \mathbb{E}\xi_{h_n,i}\right).$$

The proof make use of the Bernstein's big-blocks and small-block procedure. We divide the set $\{1,\ldots,n\}$ into $2k+1$ subsets that we note for $m=\{1,\ldots,k\}$, $l_m=(m-1)(p+q)$ and

$$s_m = (m-1)(p+q)+q.$$

We define the following sequences:

$$\begin{aligned}
S_n' &:= \sum_{m=1}^{k} \zeta_{n,m} = \sum_{m=1}^{k}\left(\sum_{i=l_m+1}^{l_m+q} Z_{n,i}\right),\\
S_n'' &:= \sum_{m=1}^{k} \zeta_{n,m}' = \sum_{m=1}^{k}\left(\sum_{i=s_m+1}^{s_m+p} Z_{n,i}\right),\\
S_n''' &:= \zeta_{n,m}'' = \sum_{i=k(p+q)+1}^{n} Z_{n,i}.
\end{aligned}$$

Then, the sum $S_n := \sum_{i=1}^{n} Z_{n,i}$ can be splitted as

$$S_n = S_n' + S_n'' + S_n'''.$$

To prove (6.3), we should show that both of $\frac{1}{\sqrt{n}}S_n''$ and $\frac{1}{\sqrt{n}}S_n'''$ are asymptotically negligible, which implies that

$$\frac{1}{n}\left[\mathbb{E}(S_n'')^2 + \mathbb{E}(S_n''')^2\right] \to 0, \tag{6.8}$$

and S' is asymptotically normal distributed as

$$\frac{1}{\sqrt{n}}\sum_{m=1}^{k} \zeta_{n,m} \xrightarrow{\mathscr{D}} \mathscr{N}(0,\Sigma_\varphi^2). \tag{6.9}$$

We first write

$$\begin{aligned}
\mathbb{E}(S_n'')^2 &= \mathbb{E}\left(\sum_{m=1}^{k}\zeta_{n,m}'\right)^2 = \operatorname{Var}\left(\sum_{m=1}^{k}\zeta_{n,m}'\right)\\
&= k\operatorname{Var}(\zeta_{n,m}') + 2\sum_{1\leqslant m<m'\leqslant k}\operatorname{Cov}(\zeta_{n,m}',\zeta_{nm'}').
\end{aligned} \tag{6.10}$$

We have

$$\begin{aligned}
\operatorname{Var}(\zeta_{n,m}') &= \operatorname{Var}\left(\sum_{i=s_m+1}^{s_m+p} Z_{n,i}\right)\\
&= p\operatorname{Var}(Z_{n,1}) + 2\sum_{s_m+1\leqslant i<j\leqslant s_m+p}\operatorname{Cov}(Z_{n,i},Z_{n,j}).
\end{aligned}$$

Clearly $Z_{n,i}$ is real valued centred random variable with

$$\operatorname{Var}(Z_{n,i}) = \frac{1}{\mathbb{E}^2\left(K_{h_n,\mathbf{x}}(\mathbf{X}_i)\right)}\operatorname{Var}(\xi_{h_n,i}).$$

Conditioning on X_i and using the assumptions (A.4) and (4.8) we deduce that

$$\begin{aligned}\mathbb{E}^2\xi_{h_n,i} &= h_n^d\mathbb{E}^2\left[(\varphi(Y_i)-m(\varphi,\mathbf{x}))K_{h_n,\mathbf{x}}(\mathbf{X}_i)\right]\\ &\leqslant h_n^d \sup_{\mathbf{u}:\|\mathbf{x}-\mathbf{u}\|\leqslant h_n}|m(\mathbf{u},\varphi)-m(\varphi,\mathbf{x})|^2\mathbb{E}^2\left(|K_{h_n,\mathbf{x}}(\mathbf{X}_i)|\right)=o(1).\end{aligned}$$

Therefore

$$\operatorname{Var}(\xi_{h_n,i}) = \mathbb{E}\xi_{h_n,1}^2+o(1).$$

Denote

$$\begin{aligned}\mathbb{E}\xi_{h_n,1}^2 &= h_n^d\mathbb{E}\left[(\varphi(Y_1)-m(\varphi,\mathbf{x}))^2K_{h_n,\mathbf{x}}^2(\mathbf{X}_1)\right]\\ &= h_n^d\mathbb{E}\left[K_{h_n,\mathbf{x}}^2(\mathbf{X}_1)\mathbb{E}((\varphi(Y_1)-m(\varphi,\mathbf{x}))^2|\mathbf{X}_1)\right].\end{aligned}$$

From the definition of the conditional variance, we have

$$\begin{aligned}\mathbb{E}\xi_{h_n,1}^2 &= h_n^d\mathbb{E}\left[K_{h_n,\mathbf{x}}^2(\mathbf{X}_1)\operatorname{Var}(\varphi(Y_1)|\mathbf{X}_1)\right]\\ &\quad + h_n^d\mathbb{E}\left[K_{h_n,\mathbf{x}}^2(\mathbf{X}_1)(m(\mathbf{X}_1,\varphi)-m(\varphi,\mathbf{x}))^2\right]\\ &= A_1+A_2.\end{aligned}$$

Write

$$\begin{aligned}A_1 &= \frac{1}{h_n^d}\int_{\mathbb{R}^d}\operatorname{Var}(\varphi(Y_1)|\mathbf{X}_1=\mathbf{u})K^2\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{u}}{h_n}\right)f_{\mathbf{X}}(\mathbf{u})d\mathbf{u}\\ &= \frac{1}{h_n^d}\int_{\mathbb{R}^d}(\sigma_\varphi^2(\mathbf{u})-\sigma_\varphi^2(\mathbf{x}))K^2\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{u}}{h_n}\right)f_{\mathbf{X}}(\mathbf{u})d\mathbf{u}\\ &\quad + \frac{\sigma_\varphi^2(\mathbf{x})}{h_n^d}\int_{\mathbb{R}^d}K^2\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{u}}{h_n}\right)f_{\mathbf{X}}(\mathbf{u})d\mathbf{u}\\ &:= A_{1.1}+A_{1.2}.\end{aligned}$$

Making use of the assumption **(A.2)**(i) and that on conditional variance given in **(A.3)**, we evaluate the term $A_{1.1}$ as follows

$$\begin{aligned}A_{1.1} &\leqslant \sup_{\mathbf{u}:\|\mathbf{x}-\mathbf{u}\|\leqslant h_n}\left|\sigma_\varphi^2(\mathbf{u})-\sigma_\varphi^2(\mathbf{x})\right|\int_{\mathbb{R}^d}K^2\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{v}\right)f_{\mathbf{X}}(\mathbf{x}+h_n\mathbf{v})d\mathbf{v}\\ &= o(1)MG_2(d)=o(1).\end{aligned}$$

Thus we have

$$A_1 = \sigma_\varphi^2(\mathbf{x})\int_{\mathbb{R}^d}K^2\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{v}\right)f_{\mathbf{X}}(\mathbf{x}+h_n\mathbf{v})d\mathbf{v}+o(1). \tag{6.11}$$

Similarly, by using the assumption **(A.4)**, we get

$$\begin{aligned}
A_2 &= h_n^d \mathbb{E}\left[(m(\mathbf{X}_1,\varphi)-m(\varphi,\mathbf{x}))^2 K_{h_n,\mathbf{x}}^2(\mathbf{X}_1)\right] \\
&\leqslant \sup_{\mathbf{u}:\|\mathbf{x}-\mathbf{u}\|\leqslant h_n} |m(\mathbf{u},\varphi)-m(\varphi,\mathbf{x})|^2 \int_{\mathbb{R}^d} K^2\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{v}\right) f_{\mathbf{X}}(\mathbf{x}+h_n\mathbf{v})d\mathbf{v} \\
&= o(1)MG_2(d) = o(1). \qquad (6.12)
\end{aligned}$$

On the other hand, from (6.2), we have

$$\mathbb{E}^2 K_{h_n,\mathbf{x}}(\mathbf{X}_1) = (f_{\mathbf{X}}(\mathbf{x})+o(1))^2 \to f_{\mathbf{X}}^2(\mathbf{x}).$$

By combining the last result with (6.11), (6.12) and (4.8), we deduce that, as $n\to\infty$,

$$\operatorname{Var}(Z_{n,1}) \to \frac{\sigma_\varphi^2(\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})} \int_{\mathbb{R}^d} K^2\left(\frac{\mathbf{x}}{h_n},\frac{\mathbf{x}}{h_n}+\mathbf{v}\right) d\mathbf{v} < \infty. \qquad (6.13)$$

Thus

$$p\operatorname{Var}(Z_{n,1}) = O(p). \qquad (6.14)$$

Now, by stationary, we have for some $c_n = o(p)$ tends to ∞ with n, the following:

$$\begin{aligned}
\sum_{\substack{i,j=s_m+1\\ i<j}}^{s_m+p} \operatorname{Cov}(Z_{n,i},Z_{n,j}) &= \sum_{j=2}^{p}(p-j)\operatorname{Cov}(Z_{n,1};Z_{n,j}) \\
&\leqslant p\left(\sum_{j=2}^{c_n}|\operatorname{Cov}(Z_{n,1},Z_{n,j})| + \sum_{j=c_n+1}^{p}|\operatorname{Cov}(Z_{n,1},Z_{n,j})|\right).
\end{aligned}$$

An application of Cauchy Schwartz inequality gives

$$|\operatorname{Cov}(Z_{n,1},Z_{n,j})| \leqslant \operatorname{Var}(Z_{n,1}),$$

and

$$\sum_{j=2}^{c_n}|\operatorname{Cov}(Z_{n,1},Z_{n,j})| \leqslant c_n \operatorname{Var}(Z_{n,1}). \qquad (6.15)$$

Recall that $(Z_{n,i})_{i\geqslant 0}$ is sequence of stationary centred random variables with

$$\mathbb{E}|Z_{n,i}|^2 = \mathbb{E}|Z_{n,1}|^2 = \operatorname{Var}(Z_{n,1}) < \infty.$$

The Lemma 3.1 combined with inequality (4.11), for $\delta = 2$, in the Appendix give the following result

$$\begin{aligned}
|\operatorname{Cov}(Z_{n,1},Z_{n,j})| &\leqslant 2\int_0^{\tilde{\alpha}_(j-1)} Q_{Z_{n,1}}^2(u)du \\
&= 2C_{2,2}\tilde{\alpha}^{\frac{1}{2}}(j-1). \qquad (6.16)
\end{aligned}$$

The assumption (**A.1**) and the fact that $c_n \to \infty$ with n, clearly gives

$$\sum_{j=c_n+1}^{p} |\mathrm{Cov}(Z_{n,1}, Z_{n,j})| < \frac{C}{c_n^{1+\frac{1}{d}}} \sum_{i=c_n}^{\infty} i^{1+\frac{1}{d}} \tilde{\alpha}^{\frac{1}{2}}(i) = o(1).$$

This when combined with (6.15), readily implies

$$2\sum_{j=1}^{p} |\mathrm{Cov}(Z_{n,1}, Z_{n,j})| \leqslant c_n \mathrm{Var}(Z_{n,1}) + o(1). \tag{6.17}$$

We have

$$\begin{aligned}
\mathrm{Var}(\zeta'_{n,m}) &= \mathrm{Var}\left(\sum_{i=s_m+1}^{s_m+p} Z_{n,i}\right) \\
&= p\mathrm{Var}(Z_{n,1}) + 2p\sum_{j=1}^{p} |\mathrm{Cov}(Z_{n,1}, Z_{n,j})| \\
&\leqslant p\mathrm{Var}(Z_{n,1}) + 2p^2\left(\mathrm{Var}(Z_{n,1}) + o(1)\right) \\
&= O(p^2).
\end{aligned} \tag{6.18}$$

Note that by setting $m' = m+1$, we get $s_{m'} = s_m + (p+q)$. In a similar way as in (6.16), by choosing $c_n = p$, one may bound the right part of (6.10) by

$$\begin{aligned}
&\sum_{1\leqslant m<m'\leqslant k} \mathrm{Cov}(\zeta'_{n,m}, \zeta'_{nm'}) \\
&= \sum_{1\leqslant m\leqslant k} \mathrm{Cov}\left(\sum_{i=s_m+1}^{s_m+p} Z_{n,i}, \sum_{j=s_m+1}^{s_m+p} Z_{n,j+p+q}\right) \\
&\leqslant \sum_{\substack{j=p+2p+1 \\ |j-i|>p}}^{(k-1)(p+q)+2q} |\mathrm{Cov}(Z_{n,i}, Z_{n,j})| \\
&\leqslant \sum_{j=p+1}^{\infty} |\mathrm{Cov}(Z_{n,1}, Z_{n,j})| = o(1).
\end{aligned} \tag{6.19}$$

Combining the results (6.18) and (6.19), we obtain

$$\mathbb{E}(S''_n)^2 = O(p^2). \tag{6.20}$$

Notice that

$$n - k(p+q) = (p+q)\left(\frac{n}{(p+q)} - k\right) \leqslant 2q.$$

Following the similar steps of the proof of (6.18) for $p := q$, we get

$$\begin{aligned}
\mathbb{E}(S'''_n)^2 &= \mathrm{Var}\zeta''_{n,m} \\
&= \sum_{i=k(p+q)+1}^{n} \mathrm{Var}Z_{n,i} + 2\sum_{k(p+q)+1\leqslant i<j\leqslant n} \mathrm{Cov}(Z_{n,i}, Z_{n,j})
\end{aligned}$$

$$\begin{aligned}
&\leqslant (n-k(p+q))\left[\operatorname{Var} Z_{n,i}+2\sum_{j=2}^{n-k(p+q)}\operatorname{Cov}(Z_{n,1},Z_{n,j})\right]\\
&\leqslant 2q\left[\operatorname{Var} Z_{n,i}+2\left(\sum_{j=1}^{c_n''}|\operatorname{Cov}(Z_{n,1},Z_{n,j})|+\sum_{j=c_n''+1}^{q}|\operatorname{Cov}(Z_{n,1},Z_{n,j})|\right)\right]\\
&=O(q^2). \qquad (6.21)
\end{aligned}$$

This when combined with (6.20) and (6.2), gives

$$\frac{1}{n\mathbb{E}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_i)\right)}[\mathbb{E}(S_n'')^2+\mathbb{E}(S_n''')^2]=O\left(\frac{p^2}{n}+\frac{q^2}{n}\right)\to 0.$$

Now, the proof of the asymptotic normality of S' in (6.9) is based on two steps. Let us begin by checking the Lindeberg-Feller conditions. More precisely, we verify

$$\lim_{n\to\infty}\frac{1}{n\mathbb{E}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_i)\right)}\sum_{m=1}^{k}\mathbb{E}\zeta_{n,m}^2=\Sigma_\varphi^2, \qquad (6.22)$$

and

$$\frac{1}{n\mathbb{E}\left(K_{h_n,\mathbf{x}}(\mathbf{X}_i)\right)}\sum_{m=1}^{k}\mathbb{E}\left(\zeta_{n,m}^2\mathbb{1}_{|\zeta_{n,m}|>\varepsilon\Sigma_\varphi\sqrt{n}}\right)\to 0. \qquad (6.23)$$

Recall that, when $n\to\infty$, $\frac{kq}{n}\to 1$. The similar calculus as that for the proof of (6.19) for $p:=q$ yield

$$\begin{aligned}
\frac{1}{n}\sum_{m=1}^{k}\mathbb{E}\zeta_{n,m}^2&=\sum_{m=1}^{k}\operatorname{Var}(\zeta_{n,m})\\
&=\frac{1}{n}\sum_{m=1}^{k}\left(\sum_{i=l_m+1}^{l_m+q}\operatorname{Var}(Z_{n,i})+\underbrace{2\sum_{i=l_m+1}^{l_m+q}\sum_{j=l_{m'}+1}^{l_m'+q}\operatorname{Cov}(Z_{n,i},Z_{n,j})}_{o(1)}\right)\\
&\to\operatorname{Var}(Z_{n,1}).
\end{aligned}$$

Put, for $\mathbf{x}\in[-2L,2L]^d$,

$$\mathbf{x}_j=\left(\frac{\lfloor 2^jx_1\rfloor}{2^j},\ldots,\frac{\lfloor 2^jx_d\rfloor}{2^j}\right),$$

where $[\mathbf{x}]$ denotes the integer part of $\mathbf{x}$ and $[\mathbf{x}]\leqslant\mathbf{x}\leqslant[\mathbf{x}]+1$. Thus, by the fact that

$$K(\mathbf{u},\mathbf{v})=K(\mathbf{u}+\mathbf{k},\mathbf{v}+\mathbf{k}),\ \text{for }\mathbf{k}\in\mathbb{Z}^d$$

we deduce from (6.13) the following:

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\sum_{m=1}^{k}\mathbb{E}\zeta_{n,m}^{2} &= \lim_{n\to\infty}\frac{1}{n}\frac{\sigma_{\varphi}^{2}(\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})}\int_{\mathbb{R}^d}K^2\left(\frac{\mathbf{x}}{h_n}-\frac{\mathbf{x}_j}{h_n},\frac{\mathbf{x}}{h_n}-\frac{\mathbf{x}_j}{h_n}+\mathbf{v}\right)d\mathbf{v}\\
&= \frac{\sigma_{\varphi}^{2}(\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})}\int_{\mathbb{R}^d}K^2\left(\mathbf{0},\mathbf{v}\right)d\mathbf{v}\\
&= \frac{\sigma_{\varphi}^{2}(\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})}\int_{\mathbb{R}^d}\left(\sum_{\mathrm{k}\in\mathbb{Z}^d}\phi(\mathbf{0}+\mathrm{k})\phi(\mathbf{u}+\mathrm{k})\right)^2 d\mathbf{u}\\
&:= \Sigma_{\varphi}^{2}.
\end{aligned}$$

Making use of Hölder (for $\wp=\frac{3}{2}$) and Markov (of order $2\wp$) inequalities, we get the following bound

$$\begin{aligned}
&\frac{1}{n\mathbb{E}^2K_{h_n,\mathbf{x}}(\mathbf{X}_1)}\sum_{m=1}^{k}\mathbb{E}\left(\zeta_{n,m}^{2}\mathbb{1}_{|\zeta_{n,m}|>\varepsilon\Sigma_{\varphi}^{2}\sqrt{n\mathbb{E}^2K_{h_n,\mathbf{x}}(\mathbf{X}_1)}}\right)\\
&\leqslant \frac{1}{n\mathbb{E}^2K_{h_n,\mathbf{x}}(\mathbf{X}_1)}\sum_{m=1}^{k}\left[\mathbb{E}(|\zeta_{n,m}|^{2\wp})^{\frac{1}{\wp}}\mathbb{P}^{1-\frac{1}{\wp}}\left(|\zeta_{n,m}|>\varepsilon\Sigma_{\varphi}^{2}\sqrt{n}\right)\right]\\
&\leqslant C\frac{1}{n\mathbb{E}^2K_{h_n,\mathbf{x}}(\mathbf{X}_1)}\sum_{m=1}^{k}\left[\left(\mathbb{E}|\zeta_{n,m}|^3\right)^{\frac{2}{3}}\left(\frac{\mathbb{E}|\zeta_{n,m}|^3}{\left(\varepsilon\Sigma_{\varphi}^{2}\sqrt{n\mathbb{E}^2K_{h_n,\mathbf{x}}(\mathbf{X}_1)}\right)^3}\right)^{\frac{1}{3}}\right]\\
&\leqslant \frac{1}{(n\mathbb{E}^2K_{h_n,\mathbf{x}}(\mathbf{X}_1))^{\frac{3}{2}}}\sum_{m=1}^{k}\mathbb{E}|\zeta_{n,m}|^3.
\end{aligned}$$

On the other hand, recall that $|x-y|^a\leqslant(|x|+|y|)^a=\sum_{k=0}^{a}C_a^k|x|^k|y|^{a-k}$. Thus by the definition of $Z_{n,i}$, the equation (4.7), and if we take condition (3.7), we obviously infer, for any $i\in\mathbb{N}$, that

$$\begin{aligned}
\mathbb{E}|Z_{n,i}|^3 &< \sum_{k=0}^{3}C_3^k h_n^{\frac{3}{2}}\mathbb{E}\left|(\varphi(Y_1)-m(\varphi,\mathbf{x}))K_{h_n,\mathbf{x}}(\mathbf{X}_1)\right|^3\\
&< 3!C_{d+1}^{3}h_n^{\frac{3}{2}}\mathbb{E}\left|\varphi(Y_1)\right|^3<\infty.
\end{aligned}$$

The use of Lemma 7.1 in the Appendix, for the order $p=3$ along with (4.11) gives the following:

$$\begin{aligned}
\mathbb{E}|\zeta_{n,m}|^3=\mathbb{E}\left|\sum_{i=l_m+1}^{l_m+q}Z_{n,i}\right|^3 &\leqslant (6q)^{\frac{3}{2}}\int_0^1\left(\sum_{r=0}^{q}\mathbb{1}_{u\leqslant\tilde{\alpha}(r)}\right)^{3/2}Q_{Z_{n,j}}^{3}(u)du\\
&\leqslant \frac{3}{2}(6q)^{\frac{3}{2}}\int_0^1\sum_{r=0}^{q}(r+1)^{\frac{1}{2}}\mathbb{1}_{u<\tilde{\alpha}(r)}Q_{Z_{n,j}}^{3}(u)du
\end{aligned}$$

$$\leqslant \frac{3}{2}(6q)^{\frac{3}{2}} \sum_{r=0}^{q} \left[(r+1)^{\frac{1}{2}} \int_0^{\tilde{\alpha}(r)} Q^3_{Z_{n,j}}(u)du \right]$$
$$\leqslant \frac{3}{2}(6q)^{\frac{3}{2}} \sum_{r=0}^{q} (r+1)^{\frac{1}{2}} \tilde{\alpha}(r)^{1-\frac{3}{3+\delta}}.$$

The later result when combined with equation (6.2) and the fact that $\frac{kq}{n} \to 1$, we have

$$\frac{1}{n\mathbb{E}^2 K_{h_n,\mathbf{x}}(\mathbf{X}_1)} \sum_{m=1}^{k} \mathbb{E}\left(\zeta^2_{n,m} \mathbb{1}_{|\zeta_{n,m}|>\varepsilon\Sigma^2_{\varphi}\sqrt{n\mathbb{E}^2 K_{h_n,\mathbf{x}}(\mathbf{X}_1)}} \right)$$
$$\leqslant C\left(\frac{q}{n}\right)^{\frac{3}{2}} \sum_{r=0}^{q} (r+1)^{\frac{1}{2}} \tilde{\alpha}(r)^{\frac{\delta}{3+\delta}}.$$

Write, for $c_n < q$

$$\left(\frac{q}{n}\right)^{\frac{3}{2}} \sum_{r=0}^{q} (r+1)^{\frac{1}{2}} \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r)$$
$$= \left(\frac{q}{n}\right)^{\frac{3}{2}} \left[\sum_{r=0}^{[c_n]} (r+1)^{\frac{1}{2}} \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r) + \sum_{r=[c_n]+1}^{q} (r+1)^{\frac{1}{2}} \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r) \right]. \tag{6.24}$$

Put $B_r := \left(\frac{q}{n}\right)^{\frac{3}{2}} \sum_{r=0}^{[c_n]} (r+1)^{\frac{1}{2}} < \left(\frac{q}{n}\right)^{\frac{3}{2}} (c_n+1)^{\frac{3}{2}}$. Using Abel transformation process, we may write

$$\left(\frac{q}{n}\right)^{\frac{3}{2}} \sum_{r=0}^{[c_n]} (r+1)^{\frac{1}{2}} \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r) = \sum_{r=0}^{[c_n]-1} \left[\tilde{\alpha}^{\frac{\delta}{\delta+3}}(r) - \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r+1) \right] B_r + B_{[c_n]} \tilde{\alpha}^{\frac{\delta}{\delta+3}}([c_n]).$$

Since $\tilde{\alpha}(r)$ is a decreasing function with limit goes to 0 when $r \to \infty$, we have

$$\sum_{r=0}^{\infty} \left[\tilde{\alpha}^{\frac{\delta}{\delta+3}}(r) - \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r+1) \right] < \infty.$$

By choosing $c_n := q-1$, we get $B_r < \left(\frac{q^2}{n}\right)^{\frac{3}{2}} = o(1)$ as $n \to \infty$. We deduce that

$$\left(\frac{q^2}{n}\right)^{\frac{3}{2}} \sum_{r=0}^{[q-1]} (r+1)^{\frac{1}{2}} \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r) = o(1) \text{ as } n \to \infty.$$

Now we control the right term of (6.24).

$$\left(\frac{q^2}{n}\right)^{\frac{3}{2}} \sum_{r=[q-1]}^{q} (r+1)^{\frac{1}{2}} \tilde{\alpha}^{\frac{\delta}{\delta+3}}(r) \leqslant \left(\frac{q^2}{n}\right)^{\frac{3}{2}} \tilde{\alpha}^{\frac{\delta}{\delta+3}}([q-1]) = o(1) \text{ as } n \to \infty.$$

Thus (6.23) is verified. To complete our proof of (6.9), all what remains is to check that

$$\left|\mathbb{E}\exp\left(it\frac{\sum_{m=1}^{k}\zeta_{n,m}}{\sqrt{n}}\right)-\prod_{m=1}^{k}\mathbb{E}\exp\left(it\frac{\zeta_{n,m}}{\sqrt{n}}\right)\right|\to 0.$$

We set, for some $y\in\mathbb{R}$,

$$\rho(y)=\exp\left(it\frac{y}{\sqrt{n}}\right).$$

One can see that, for any variable $\mathbf{x}\in\mathbb{R}^d$, $\rho(\cdot)$ is an exponential holomorphic function, which implies that's Lipschitz function with

$$\operatorname{Lip}\exp\left(it\frac{y}{\sqrt{n}}\right)\leqslant\frac{|t|}{\sqrt{n}},$$

and uniformly bounded by 1. We write

$$\begin{aligned}
&\left|\mathbb{E}\prod_{m=1}^{k}\rho\left(\zeta_{n,m}\right)-\prod_{m=1}^{k}\mathbb{E}\rho\left(\zeta_{n,m}\right)\right|\\
&\leqslant\left|\mathbb{E}\prod_{m=1}^{k}\rho\left(\zeta_{n,m}\right)-\mathbb{E}\rho\left(\zeta_{n,k}\right)\mathbb{E}\prod_{m=1}^{k-1}\rho\left(\zeta_{n,m}\right)\right|+\left|\mathbb{E}\prod_{m=1}^{k-1}\rho\left(\zeta_{n,m}\right)-\prod_{j=1}^{k-1}\mathbb{E}\rho\left(\zeta_{n,m}\right)\right|\\
&=\left|\operatorname{Cov}\left(\prod_{m=1}^{k-1}\rho\left(\zeta_{n,m}\right),\rho\left(\zeta_{n,k}\right)\right)\right|+\left|\mathbb{E}\prod_{m=1}^{k-1}\rho\left(\zeta_{n,m}\right)-\prod_{j=1}^{k-1}\mathbb{E}\rho\left(\zeta_{n,m}\right)\right|\\
&=\sum_{s=1}^{k}\left|\operatorname{Cov}\left(\prod_{m=1}^{k-s}\rho\left(\zeta_{n,m}\right),\rho\left(\zeta_{n,k-s+1}\right)\right)\right|.
\end{aligned}$$

We define, for any fixed $s<k$,

$$g:\left(\zeta_{n,1},\ldots,\zeta_{n,k-s}\right)\longrightarrow\prod_{m=1}^{k-s}\rho\left(\zeta_{n,m}\right)=\rho\left(\sum_{m=1}^{k-s}\sum_{i=l_m+1}^{l_m+q}Z_{n,i}\right).$$

Using the Proposition 3.2, for $\mathbf{X}_i=\left(\zeta_{n,1},\ldots,\zeta_{n,k-s}\right)$ and $\mathbf{X}_j=\zeta_{n,k}=\sum_{i=l_k+1}^{l_k+q}Z_{n,i}$, we get

$$\left|\operatorname{Cov}\left(g\left(\mathbf{X}_i\right),\rho\left(\zeta_{n,k}\right)\right)\right|\leqslant C\frac{|t|}{\sqrt{n}}\tilde{\alpha}^{\frac{1}{2}}(r),$$

where $r:=|l_k+1-(l_{k-1}+q)|=p+1$. We deduce that

$$\begin{aligned}
\left|\mathbb{E}\prod_{m=1}^{k}\rho\left(\zeta_{n,m}\right)-\prod_{m=1}^{k}\mathbb{E}\rho\left(\zeta_{n,m}\right)\right|&\leqslant 4k\frac{|t|}{\sqrt{n}}\tilde{\alpha}^{\frac{1}{2}}(p)\\
&\leqslant C\frac{k}{\sqrt{n}}\tilde{\alpha}^{\frac{1}{2}}(p)=o(1).
\end{aligned}$$

By that, the proof of Lemma 6.3 is completed, allowing us to conclude the result of the main Theorem 4.1. □

7. Appendix

Proof of Proposition 3.2

For Ω rich enough, we can define the random vectors $\mathbf{X}_j$ by $\mathbf{X}^*_{j,d_1} = (X^*_{j_1},\dots,X^*_{j_{d_1}})$ independent of $\mathbf{X}_i$ and distributed as $\mathbf{X}_{j,d_1} = (X_{j_1},\dots,X_{j_{d_1}})$. By stationary and since $g_2(\cdot)$ is Lipschitz continuous function in $\mathbb{R}^{d_1} \to \mathbb{R}$, we get

$$\begin{aligned}
\left|\mathrm{Cov}(g_1(\mathbf{X}_i), g_2(\mathbf{X}_j))\right| &= \left|\mathbb{E}\left(g_1(\mathbf{X}_i)(g_2(\mathbf{X}^*_j) - g_2(\mathbf{X}_j))\right)\right| \\
&\leqslant \|g_1\|_\infty \mathrm{Lip}(g_2)\mathbb{E}\|\mathbf{X}^*_j - \mathbf{X}_j\| \\
&= \|g_1\|_\infty \mathrm{Lip}(g_2)\sum_{l=1}^{d_1}\mathbb{E}|X^*_{j_l} - X_{j_l}| \\
&\leqslant \|g_1\|_\infty \mathrm{Lip}(g_2) d_1 \max_{1\leqslant l\leqslant d_1}\mathbb{E}|X^*_{j_l} - X_{j_l}|. \qquad (7.1)
\end{aligned}$$

Recall that $\mathbf{X}_j$ is a random vector of stationary integrable random variables and $\widetilde{\alpha}$-weak dependent defined on probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\mathscr{M}$ be a sub-sigma algebra of $\mathscr{F}$. For any $1 \leqslant l \leqslant d_1$, X_{j_l} a component of $\mathbf{X}_j$ is an $\widetilde{\alpha}$-weak dependence random variable. An important result belonging to $\widetilde{\alpha}$-weak dependence for such special case is proved in [45], by proceeding the coupling results of [36]'s quantile transformation (we refer to the paper by [40] for more details on coupling techniques and their applications). Define some $U_{j_l} \sim U[0,1]$, that is independent of the σ-algebras generated by $\mathscr{M}$ and X_{j_l}. The random variable $X^*_{j_l}$, is measurable with respect to $\mathscr{M} \vee \sigma(X_{j_l}, U_{j_l})$. The Theorem 2 (c) of [45], gives the following result

$$\mathbb{E}|X^*_{j_l} - X_{j_l}| \leqslant 4\int_0^{\widetilde{\alpha}(\mathscr{M}_{j_l}, X_{j_l})} Q_{|X_{j_l}|}(u)du,$$

such $\mathscr{M}_{j_l} = \sigma(X_i, i < j_l)$. For any $(i, j_l) \in \Delta(1,l,r)$ we have $\mathscr{M}_i \subset \mathscr{M}_{j_l}$ and if $\mathbb{E}|\mathbf{X}|^2 < \infty$, the definition of $\tilde{\alpha}$ coefficient and an application of Hölder and Jensen inequalities yields to

$$\begin{aligned}
\max_{1\leqslant l\leqslant d_1}\sup_{(i,j)\in\Delta(1,l,r)}\mathbb{E}|X^*_{j_l} - X_{j_l}| &\leqslant 4\left(\int_0^1 Q^2_{|X_{j_l}|}(u)du\right)^{1/2}\left(\int_0^1 \mathbb{1}_{u\leqslant\widetilde{\alpha}(r)}du\right)^{1/2} \\
&= 4\|X\|_2\widetilde{\alpha}^{\frac{1}{2}}(r) \\
&= C\widetilde{\alpha}^{\frac{1}{2}}(r). \qquad (7.2)
\end{aligned}$$

Combining results in (7.1) and (7.2), we get the upper bound of (3.2). □

LEMMA 7.1. [19, Corollary 5.3] *Let $(X)_{i\in\mathbb{N}}$ be a sequence of centered and square-integrable random variables and $\mathscr{M}_i = \sigma(X_j, 0 \leqslant j \leqslant i)$. Define $\tilde{\alpha}_i = \sup_{r\geqslant 0}\tilde{\alpha}(\mathscr{M}_r, X_{i+r})$.*

Note

$$\tilde{\alpha}_n^{-1}(u) = \sum_{i=0}^{n}\mathbb{1}_{u\leqslant\tilde{\alpha}_i}.$$

For any random variable X such that $Q_X \geqslant \sup_{r\geqslant 1} Q_{X_r}$. For $p \geqslant 2$, we have

$$\left\| \sum_{i=1}^{n} X_i \right\|_p \leqslant \sqrt{2pn} \left(\int_0^1 \left(\tilde{\alpha}_r^{-1}(u)\right)^{\frac{p}{2}} Q_X^p du \right)^{1/p}.$$

REMARK 7.1. We highlight that

$$\left(\tilde{\alpha}_r^{-1}(u)\right)^{\frac{p}{2}} = \sum_{r=0}^{n} \left((r+1)^{\frac{p}{2}} - r^{\frac{p}{2}} \right) \mathbb{1}_{u \leqslant \tilde{\alpha}(r)} \leqslant \sum_{r=0}^{n} \left(1 \vee \frac{p}{2} \right) (r+1)^{\frac{p}{2}-1} \mathbb{1}_{u \leqslant \tilde{\alpha}(r)}.$$

Proof of the Remark 7.1

Recall that $(\tilde{\alpha}_r) r \geqslant 0$ is a decreasing sequence of non negative numbers. For any function $\vartheta(\cdot)$, we may write

$$\vartheta\left(\tilde{\alpha}_r^{-1}(u)\right) = \sum_{r=0}^{n} \vartheta(r+1) \mathbb{1}_{u \leqslant \tilde{\alpha}(r)},$$

and if $\vartheta(0) = 0$ we have

$$\vartheta(r+1) = \sum_{j=0}^{r} \vartheta(j+1) - \vartheta(j),$$

we refer to [18, proof of Lemma 2]. Then the last assertion follows by choosing $\vartheta(x) = x^{\frac{p}{2}}$ and by the fact that

$$(r+1)^{\frac{p}{2}} - r^{\frac{p}{2}} \leqslant \left(1 \vee \frac{p}{2}\right) (r+1)^{\frac{p}{2}-1}. \quad \square$$

Acknowledgement. The authors are indebted to the Editor-in-Chief, Associate Editor and the referee for their very valuable comments, suggestions careful reading of the article which led to a considerable improvement of the manuscript.

REFERENCES

[1] H. AKAIKE, *An approximation to the density function*, Ann. Inst. Statist. Math., Tokyo, **6** (1954), 127–132.

[2] S. ALLAOUI, S. BOUZEBDA AND J. LIU, *Multivariate wavelet estimators for weakly dependent processes: strong consistency rate*, Comm. Statist. Theory Methods, **0** (0) (2022), 1–34.

[3] D. W. K. ANDREWS, *Non-strong mixing autoregressive processes*, J. Appl. Probab., **21** (1984), 930–934.

[4] P. J. BICKEL AND P. BÜHLMANN, *A new mixing notion and functional central limit theorems for a sieve bootstrap in time series*, Bernoulli, **5** (3) (1999), 413–446.

[5] S. BOUZEBDA AND S. DIDI, *Multivariate wavelet density and regression estimators for stationary and ergodic discrete time processes: asymptotic results*, Comm. Statist. Theory Methods, **46** (3) (2017), 1367–1406.

[6] S. BOUZEBDA AND S. DIDI, *Some asymptotic properties of kernel regression estimators of the mode for stationary and ergodic continuous time processes*, Rev. Mat. Complut., **34** (3) (2021), 811–852.

[7] S. BOUZEBDA AND S. DIDI, *Some results about kernel estimators for function derivatives based on stationary and ergodic continuous time processes with applications*, Comm. Statist. Theory Methods, **51** (12) (2022), 3886–3933.

[8] S. BOUZEBDA AND B. NEMOUCHI, *Uniform consistency and uniform in bandwidth consistency for nonparametric regression estimates and conditional U-statistics involving functional data*, J. Nonparametr. Stat., **32** (2) (2020), 452–509.

[9] S. BOUZEBDA AND B. NEMOUCHI, *Weak-convergence of empirical conditional processes and conditional U-processes involving functional mixing data*, Stat. Inference Stoch. Process. To appear, (2022) pages 1–56.

[10] S. BOUZEBDA AND A. NEZZAL, *Uniform consistency and uniform in number of neighbors consistency for nonparametric regression estimates and conditional U-statistics involving functional data*, Jpn. J. Stat. Data Sci., **5** (2) (2022), 431–533.

[11] S. BOUZEBDA AND N. TAACHOUCHE, *Rates of the strong uniform consistency for the kernel-type regression function estimators with general kernels on manifolds*, Math. Methods Statist., pages 1–62, 2022.

[12] S. BOUZEBDA, S. DIDI AND L. EL HAJJ, *Multivariate wavelet density and regression estimators for stationary and ergodic continuous time processes: asymptotic results*, Math. Methods Statist., **24** (3) (2015), 163–199.

[13] S. BOUZEBDA, T. EL-HADJALI AND A. A. FERFACHE, *Uniform in bandwidth consistency of conditional U-statistics adaptive to intrinsic dimension in presence of censored data*, Sankhya A, pages 1–59, 2022.

[14] J. E. CHACÓN AND T. DUONG, *Multivariate kernel smoothing and its applications*, volume 160 of *Monographs on Statistics and Applied Probability*, CRC Press, Boca Raton, FL. 2018.

[15] C. CHESNEAU AND H. DOOSTI, *A note on the adaptive estimation of a conditional continuous-discrete multivariate density by wavelet methods*, Chin. J. Math. (N.Y.), pages Art. ID 6204874, 8, 2016.

[16] C. CHESNEAU, H. DOOSTI AND L. STONE, *Adaptive wavelet estimation of a function from an m-dependent process with possibly unbounded m*, Comm. Statist. Theory Methods, **48** (5) (2019), 1123–1135.

[17] I. DAUBECHIES, *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[18] J. DEDECKER AND P. DOUKHAN, *A new covariance inequality and applications*, Stoch Process Their Appl, **106** (1) (2002), 63–80.

[19] J. DEDECKER, P. DOUKHAN, G. LANG, J. R. LEÓN, S. LOUHICHI AND C. PRIEUR, *Weak dependence: with examples and applications*, volume 190 of *Lecture Notes in Statistics*, Springer, New York, 2007.

[20] L. DEVROYE, *A course in density estimation*, volume 14 of *Progress in Probability and Statistics*, Birkhäuser Boston, Inc., Boston, MA, 1987.

[21] L. DEVROYE AND G. LUGOSI, *Combinatorial methods in density estimation*, Springer Series in Statistics, Springer-Verlag, New York, 2001.

[22] S. DIDI AND S. BOUZEBDA, *Wavelet density and regression estimators for continuous time functional stationary and ergodic processes*, Mathematics, **10** (22) (2022).

[23] S. DIDI, A. ALHARBY AND S. BOUZEBDA, *Wavelet density and regression estimators for functional stationary and ergodic data: Discrete time*, Mathematics, **10** (19) (2022).

[24] D. L. DONOHO, I. M. JOHNSTONE, G. KERKYACHARIAN AND D. PICARD, *Density estimation by wavelet thresholding*, Ann. Statist., **24** (2) (1996), 508–539.

[25] P. DOUKHAN, *Stochastic models for time series*, volume 80 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*, Springer, Cham, 2018.

[26] P. DOUKHAN AND S. LOUHICHI, *A new weak dependence condition and applications to moment inequalities*, Stochastic Process. Appl., **84** (2) (1999), 313–342.

[27] P. DOUKHAN AND S. LOUHICHI, *Functional estimation of a density under a new weak dependence condition*, Scand. J. Statist., **28** (2) (2001), 325–341.

[28] P. P. B. EGGERMONT AND V. N. LARICCIA, *Maximum penalized likelihood estimation. Volume II. Regression*, Springer Series in Statistics. Springer, Dordrecht, 2009.

[29] J. C. ESCANCIANO, *Uniform rates for kernel estimators of weakly dependent data*, 2020.

[30] E. GINÉ AND R. NICKL, *Uniform limit theorems for wavelet density estimators*, Ann. Probab., **37** (4) (2009), 1605–1646.

[31] E. GINÉ AND R. NICKL, *Mathematical foundations of infinite-dimensional statistical models*, Cambridge Series in Statistical and Probabilistic Mathematics, [40], Cambridge University Press, New York, 2016.

[32] L. GYÖRFI, M. KOHLER, A. KRZYŻAK AND H. WALK, *A distribution-free theory of nonparametric regression*, Springer Series in Statistics, Springer-Verlag, New York, 2002.

[33] P. HALL AND J. HOROWITZ, *A simple bootstrap method for constructing nonparametric confidence bands for functions*, Ann. Statist., **41** (4) (2013), 1892–1921.

[34] P. HALL AND S. PENEV, *Cross-validation for choosing resolution level for nonlinear wavelet curve estimators*, Bernoulli, **7** (2) (2001), 317–341.

[35] W. HÄRDLE, G. KERKYACHARIAN, D. PICARD AND A. TSYBAKOV, *Wavelets, approximation, and statistical applications*, volume 129 of *Lecture Notes in Statistics*, Springer-Verlag, New York, 1998.

[36] P. MAJOR, *On the invariance principle for sums of independent identically distributed random variables*, J. Multivariate Anal., **8** (4) (1978), 487–517.

[37] E. MASRY, *Probability density estimation from dependent observations using wavelets orthonormal bases*, Statist. Probab. Lett., **21** (3) (1994), 181–194.

[38] E. MASRY, *Multivariate probability density estimation by wavelet methods: strong consistency and rates for stationary time series*, Stochastic Process. Appl., **67** (2) (1997), 177–193.

[39] E. MASRY, *Wavelet-based estimation of multivariate regression functions in Besov spaces*, J. Nonparametr. Statist., **12** (2) (2000), 283–308.

[40] F. MERLEVÈDE AND M. PELIGRAD, *On the coupling of dependent random variables and applications*, In *Empirical process techniques for dependent data*, pages 171–193, Birkhäuser Boston, Boston, MA, 2002.

[41] Y. MEYER, *Wavelets and operators*, volume 37 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, 1992.

[42] E. A. NADARAJA, *On a regression estimate*, Teor. Verojatnost. i Primenen., **9** (1964), 157–159.

[43] E. A. NADARAJA, *Nonparametric estimation of probability densities and regression curves*, volume 20 of *Mathematics and its Applications (Soviet Series)*, Kluwer Academic Publishers Group, Dordrecht, 1989.

[44] E. PARZEN, *On estimation of a probability density function and mode*, Ann. Math. Statist., **33** (1962), 1065–1076.

[45] M. PELIGRAD, *Some remarks on coupling of dependent random variables*, Statist. Probab. Lett., **60** (2) (2002), 201–209.

[46] N. RAGACHE AND O. WINTENBERGER, *Convergence rates for density estimators of weakly dependent time series*, In *Dependence in probability and statistics*, volume 187 of *Lect. Notes Stat.*, pages 349–372, Springer, New York, 2006.

[47] E. RIO, *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*, Springer-Verlag, Berlin, 2000.

[48] M. ROSENBLATT, *A central limit theorem and a strong mixing condition*, Proc. Nat. Acad. Sci. U.S.A., **42** (1956a), 43–47.

[49] M. ROSENBLATT, *A central limit theorem and a strong mixing condition*, Proc. Nat. Acad. Sci. U.S.A., **42** (1956b), 43–47.

[50] D. W. SCOTT, *Multivariate density estimation. Theory, practice, and visualization*, Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2015.

[51] B. W. SILVERMAN, *Density estimation for statistics and data analysis*, Monographs on Statistics and Applied Probability, Chapman & Hall, London, 1986.

[52] B. VIDAKOVIC, *Statistical modeling by wavelets*, Wiley Series in Probability and Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.

[53] M. P. WAND AND M. C. JONES, *Kernel smoothing*, volume 60 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, Ltd., London, 1995.

[54] L. WASSERMAN, *All of nonparametric statistics*, Springer Texts Stat. New York, NY: Springer, 2006.

[55] G. S. WATSON, *Smooth regression analysis*, Sankhyā Ser. A, **26** (1964), 359–372.

[56] X. ZHOU, *Uniform convergence rates for wavelet curve estimation in sup-norm loss*, J. Comput. Appl. Math., **400** (2022), Paper No. 113752, 17.

[57] Y.-Y. ZOU, H.-Y. LIANG AND J.-J. ZHANG, *Nonlinear wavelet density estimation with data missing at random when covariates are present*, Metrika, **78** (8) (2015), 967–995.

(Received January 3, 2022)

Soumaya Allaoui
School of Mathematics and Statistics
Huazhong University of Science and Technology
Wuhan, China
e-mail: soum_allaoui@hust.edu.cn

Salim Bouzebda
Université de technologie de Compiègne
LMAC (Laboratory of Applied Mathematics of Compiègne)
CS 60 319–60 203 Compiègne Cedex
e-mail: salim.bouzebda@utc.com

Jicheng Liu
School of Mathematics and Statistics
Huazhong University of Science and Technology
Wuhan, China
e-mail: jclui@hust.edu.cn