

SOME CHARACTERISTICS OF THE CONDITIONAL SET-INDEXED EMPIRICAL PROCESS INVOLVING FUNCTIONAL ERGODIC DATA

YOUSOUF SOUDDI^{1,a}, FETHI MADANI^{2,b} AND SALIM BOUZEBDA^{3,c}

¹Laboratory of Stochastic Models, Statistics and Applications, University of Saida-Dr. Moulay Tahar,
P.O. Box 138, EN-NASR, 20000, Algeria.

^aE-mail: souddiyoucef@yahoo.fr

²Laboratory of Stochastic Models, Statistics and Applications, University of Saida-Dr. Moulay Tahar,
P.O. Box 138, EN-NASR, 20000, Algeria.

^bE-mail: fethi.madani@univ-saida.dz

³Alliance Sorbonne Université, Université de Technologie de Compiègne, L.M.A.C., Compiègne, France.
^cE-mail: bouzebda@gmail.com

Abstract

The purpose of this paper is to establish the invariance principle for the conditional set-indexed empirical process formed by functional ergodic random variables. The limit theorems, discussed in this paper, are key tools for many further developments in functional data analysis involving empirical process techniques. These results are proved under some standard structural conditions on the Vapnik-Chervonenkis classes of functions and some mild conditions on the model.

1. Introduction

The theory of empirical process is branch of statistics and play fundamental role in its various applications especially important in estimation theory there has been a great deal research works. The asymptotic properties of empirical processes indexed by functions have been intensively studied during the past decades (see, e.g., [55] or [26] for self-contained, comprehensive books on the topic with various statistical applications). Many authors have studied it in the last century in finite framework, so that its developed rapidly due to its role in solving problems of statistics, modulo measurability,

Received December 2, 2021.

AMS Subject Classification: 31B30, 35J35.

Key words and phrases: Conditional distribution, Nadaraya-Watson regression estimator, empirical process, ergodic, functional data, semi-metric space, covering number, small ball probability.

the classes \mathcal{C} of sets for which the Glivenko-Cantelli theorem holds characterized by [56] in the setting of independent variables and in this framework many results had been obtained we cite [24], [30], [41], [44] and [5]. Empirical processes based on dependent data have been studied under various mixing conditions, for example, [?] established the asymptotic normality when the sequences are ϕ -mixing, in these lines of research in different type of mixing, we may cite [27], [43] and [59]. However, a bracketing condition under strong mixing was stated by [1]. [23] studied the function-indexed empirical process for β -mixing sequences, where [3] was given results the case of Gaussian long-range dependent random vectors, [45] have established uniform convergence and asymptotic normality of set-indexed conditional empirical process in a strictly stationary and strong mixing framework and derived the Bahadur Kiefer approximations of conditional quantile in this framework [46] extended the work of [45]. On the other hand, the modelization of functional variables that taking values in infinite dimensional spaces had received a lot of attention in the last few years, there are an increasing number of situation coming from different fields of applied sciences (environment, chemometrics, biometrics, medicine, econometrics,...) in which the collected data are curves, the study of statistical models adapted to such type of infinite dimensional data has been the subject of several works in the recent statistical literature good overviews about this literature can be found in [48], [7], [49], [28], [8], [52], [36], [60], [6], [38] and [2] and hundreds of papers and books have been published in this framework last decade.

However, there are a few results for the empirical process considered functional framework, we may refer for recent references to [14, 15, 16], [9]. [19] obtained several very useful results for set-indexed conditional empirical processes in functional setting the strong mixing dependence. Notice that mixing is some kind of asymptotic independence assumption which is commonly used for seeking simplicity but which can be unrealistic in situations where there is strong dependence between the data. Extending non-parametric functional ideas to general dependence structure is a rather underdeveloped field, the ergodic framework avoids the widely used strong mixing condition and its variants to measure the dependency which go far beyond the invariance principle that is the basic motivation of the paper. The general framework of ergodic functional data has been initiated by [40] who stated consistencies with rates together with the asymptotic normality of the regression function estimate, for recent paper on the subject we refer

to [9], where the authors extended the last reference to a more general framework. For reader convenience, we introduce some details defining the ergodic property of processes and its link with the mixing one. Let $\{X_n, n \in \mathbb{Z}\}$ be a stationary sequence. Consider the backward field $\mathcal{A}_n = \sigma(X_k : k \leq n)$ and the forward field $\mathcal{B}_m = \sigma(X_k : k \geq m)$. The sequence is strongly mixing if, as $n \rightarrow \infty$,

$$\sup_{A \in \mathcal{A}_0, B \in \mathcal{B}_n} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| = \alpha(n) \rightarrow 0.$$

The sequence is ergodic if

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \mathbb{P}(A \cap \tau^{-k}B) - \mathbb{P}(A)\mathbb{P}(B) \right| = 0,$$

where τ is the time-evolution or shift transformation. The naming of strong mixing in the above definition is more stringent than what is ordinarily referred (when using the vocabulary of measure preserving dynamical systems) as strong mixing, namely to that $\lim_{n \rightarrow \infty} \mathbb{P}(A \cap \tau^{-n}B) = \mathbb{P}(A)\mathbb{P}(B)$ for any two measurable sets A, B , see, for instance [50]. Hence, strong mixing implies ergodicity, whereas the inverse is not always true (see e.g. Remark 2.6 in page 50 in connection with Proposition 2.8 in page 51 in [20]). Some motivations to consider ergodic dependence structure in the data rather than a mixing one are discussed in [40, 17, 11, 10, 12] where details on the definition of ergodic property of processes together with illustrating examples of such processes are also given.

The aim of the present paper is to extend asymptotic results for set-indexed conditional empirical processes to the context of functional ergodic data. We establish uniform convergence and asymptotic normality when the observations are assumed to be ergodic in nature taking their values in semi-metric space. This paper responds to a problem that has not been studied systematically up to the present

The remainder of this paper is organized as follows. Section 2, we present the notation and definitions together with the conditional empirical process. Section 3, we give our main results. We discuss the bandwidth selection procedure in Section 3.1. An application of our main result to the test of the conditional independence is given in Section 4. Some concluding remarks and possible future developments are relegated to Section 5. To prevent from

interrupting the flow of the presentation, all proofs are gathered in Section 6. Some examples are collected in Section 7.

2. The Set Indexed Conditional Empirical Process

For the sake of clarity, introduce some details defining the ergodic property of processes. Taking a measurable space (S, \mathcal{J}) denote by $S^{\mathbb{N}}$ the space of all functions $s : \mathbb{N} \rightarrow S$. If s_j is the value the function s takes at $j \in \mathbb{N}$, define H_j as the j -th coordinate map, i.e $H_j(s) = s_j$ and consider $H_j^{-1}(\mathcal{J}), j \in \mathbb{N}$ a random process $Z = \{Z_j : j \in \mathbb{N}\}$ can be considered as random variable defined on probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and taking values in $(S^{\mathbb{N}}, \mathcal{J}^{\mathbb{N}})$. Now a set $B \in \mathcal{F}$ is called invariant if there exists some set $\mathcal{A} \in \mathcal{J}^{\mathbb{N}}$ such that $B = \{(Z_n, Z_{n+1}, \dots) \in \mathcal{A}\}$ is true for any $n \geq 1$. The process Z is then said ergodic whenever, for any invariant set B , we have $\mathbb{P}(B) = 0$ or $\mathbb{P}(\Omega | B) = 0$. It is well known from the ergodic theorem that, for a stationary ergodic process Z , we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n Z_t = \mathbb{E}(Z_1) \text{ almost surely.} \tag{1}$$

Therefore, the ergodic property in our setting is formulated on the basis of the statement (1). We consider a sample of random elements $(X_1, Y_1), \dots, (X_n, Y_n)$ copies of (X, Y) that takes its value in a space $\mathcal{E} \times \mathbb{R}^d$. The functional space \mathcal{E} is equipped with a semi-metric $d_{\mathcal{E}}(\cdot, \cdot)$. We aim to study the links between X and Y , by estimating functional operators associated to the conditional distribution of Y given X such as the regression operator, for some measurable set C in a class of sets \mathcal{C} ,

$$\mathbb{G}(C | x) = \mathbb{E}(\mathbb{1}_{\{Y \in C\}} | X = x).$$

This regression relationship suggests to consider the following Nadaraya Watson-type ([42] and [57]) conditional empirical distribution:

$$\mathbb{G}_n(C, x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{Y_i \in C\}} K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}, \tag{2}$$

where $K(\cdot)$ is a real-valued kernel function from $[0, \infty)$ into $[0, \infty)$ and h_n is a smoothing parameter satisfying $h_n \rightarrow 0$ as $n \rightarrow \infty$, C is a measurable set, and $x \in \mathcal{E}$. By choosing $C = (-\infty, z]$, $z \in \mathbb{R}^d$, it reduces to the conditional empirical distribution function $F_n(z|x) = \mathbb{G}_n((-\infty, z], x)$, refer to [53], [54], [37]. However, the corresponding class $\mathcal{C} = \{(-\infty, z], z \in \mathbb{R}^d\}$. Concerning the semi-metric topology defined on \mathcal{E} , we will use the notation

$$B(x, t) = \{x_1 \in \mathcal{E} : d_{\mathcal{E}}(x_1, x) \leq t\},$$

for the ball in \mathcal{E} with center x and radius t , usually called in the literature the small ball probability function when t is decreasing to zero. This notion plays a major role both from theoretical and practical points of view, because the notion of ball is strongly linked with the semi-metric $d(\cdot, \cdot)$, the choice of this semi-metric will become an important stage when the data is taking its values in some infinite dimensional space. Indeed, in many examples, the small ball probability function can be written approximately as the product of two independent functions in terms of x and h , as in the following examples, which can be found in Proposition 1 of [29]:

- (1) $\phi(h_n) = Ch_n^v$ for some $v > 0$ with $\tau_0(s) = s^v$;
- (2) $\phi(h_n) = Ch_n^v \exp(-Ch_n^{-p})$ for some $v > 0$ and $p > 0$ with $\tau_0(s)$ is the Dirac's function;
- (3) $\phi(h_n) = C |\ln(h_n)|^{-1}$ with $\tau_0(s) = \mathbb{1}_{]0,1]}(s)$ the indicator function in $]0, 1]$.

Let \mathcal{F}_i be the σ -field generated by $((X_1, Y_1), \dots, (X_i, Y_i))$ and \mathfrak{G}_i that generated by $((X_1, Y_1), \dots, (X_i, Y_i), X_{i+1})$. Let $B(x, u)$ be a ball centered at $x \in \mathcal{E}$ with radius u . Let $D_i = d(x, X_i)$ so that D_i is a nonnegative real-valued random variables. Working on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let

$$F_x(u) = \mathbb{P}(D_i \leq u) = \mathbb{P}(X_i \in B(x, u)),$$

and $F_x^{\mathcal{F}_{i-1}} = \mathbb{P}(X_i \in B(x, u) \mid \mathcal{F}_{i-1})$ be the distribution function and the conditional distribution function, given the σ -field \mathcal{F}_{i-1} of $(D_i)_{i \geq 1}$ respectively. Denote by $o_{a.s.}(u)$ a real random function l such that $l(u)/u$ converges to zero almost surely as $u \rightarrow 0$. Similarly define $\mathcal{O}_{a.s.}(u)$ as a real random function l such that $l(u)/u$ is almost surely bounded.

Throughout the sequel, we assume tacitly that the sequence of random elements $\{(X_i, Y_i), i = 1, \dots, n\}$ is ergodic.

2.1. Assumptions and notation

Throughout this paper x is a fixed element of the functional space \mathcal{E} . We define the metric entropy with inclusion which provides a measure of richness (or complexity) of the class of sets \mathcal{C} . For each $\varepsilon > 0$, the covering number is defined as :

$$\begin{aligned} \mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}(\cdot | x)) &= \inf\{n \in \mathbb{N} : \exists C_1, \dots, C_n \in \mathcal{C} \text{ such that } \forall C \in \mathcal{C} \exists 1 \leq i, j \leq n \\ &\text{with } C_i \subset C \subset C_j \text{ and } \mathbb{G}(C_j \setminus C_i | x) < \varepsilon\}, \end{aligned}$$

the quantity $\log(\mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}(\cdot | x)))$ is called metric entropy with inclusion of \mathcal{C} with respect to $\mathbb{G}(\cdot | x)$. Estimates for such covering numbers are known for many classes; see, e.g., [25]. We will often assume below that either $\log \mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}(\cdot | x))$ or $\mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}(\cdot | x))$ behave like powers of ε^{-1} . We say that the condition (R_γ) holds if

$$\log \mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}(\cdot | x)) \leq H_\gamma(\varepsilon), \text{ for all } \varepsilon > 0, \tag{3}$$

where

$$H_\gamma(\varepsilon) = \begin{cases} \log(A\varepsilon) & \text{if } \gamma = 0, \\ A\varepsilon^{-\gamma} & \text{if } \gamma > 0, \end{cases}$$

for some constants $A, r > 0$. As in [45], it is worth noticing that the condition (3), $\gamma = 0$, holds for intervals, rectangles, balls, ellipsoids, and for classes which are constructed from the above by performing set operations union, intersection and complement finitely many times. The classes of convex sets in \mathbb{R}^d ($d \geq 2$) fulfill the condition (3), $\gamma = (d - 1)/2$. This and other classes of sets satisfying (3) with $\gamma > 0$, can be found in [25].

Example 1 ([18]). The set \mathcal{C} all indicator functions $\mathbb{1}_{(\infty, t]}$ of cells in \mathbb{R} satisfies

$$\mathcal{N}\left(\varepsilon, \mathcal{C}, d_\gamma^{(2)}\right) \leq \frac{2}{\varepsilon^2}$$

for any probability measure γ and $\epsilon \leq 1$. Notice that

$$\int_0^1 \sqrt{\log\left(\frac{1}{\epsilon}\right)} d\epsilon \leq \int_0^\infty u^{1/2} \exp(-u) du \leq 1$$

For more details and discussion on this example refer to example 2.5.4 in [55].

We give now further notation. For $j \geq 1$ set

$$M_j = K^j(1) - \int_0^1 (K^j)'(u) \tau_0(u) du.$$

In this section, we establish the weak convergence of the process $\{\tilde{v}_n(C, x) : C \in \mathcal{C}\}$ defined by

$$\tilde{v}_n(C, x) := \sqrt{n\phi(h_n)} (\mathbb{G}_n(C, x) - \mathbb{E}\mathbb{G}_n(C, x)). \tag{4}$$

In our analysis, we will make use of the following assumptions.

(H1) For $x \in \mathcal{E}$, there exists a sequence of nonnegative bounded random functionals $(f_{i,1})_{i \geq 1}$, a sequence of random functions $(g_{i,x})_{i \geq 1}$ a deterministic nonnegative bounded functional f_1 and a nonnegative real function ϕ where $\phi(h_n) \rightarrow 0$ as $h \rightarrow 0$ such that

- (i) $F_x(u) = \phi(u)f_1(x) + o(\phi(u))$ as $u \rightarrow 0$.
- (ii) For any $i \in \mathbb{N}$, $F_x^{\tilde{v}_{i-1}}(u) = \phi(u)f_{i,1}(x) + g_{i,x}(u) = o_{a.s.}(\phi(u))$ as $u \rightarrow 0$. $g_{i,x}(u)/\phi(u)$ almost surely bounded and $n^{-1} \sum_{t=1}^n g_{i,x}^j(u) = o_{a.s.}(\phi^j(u))$ as $n \rightarrow \infty, j = 1, 2$.
- (iii) $n^{-1} \sum_{t=1}^n f_{i,1}^j(x) \rightarrow f_1^j(x)$ almost surely as $n \rightarrow \infty$, for $j = 1, 2$.
- (iv) There exists nondecreasing bounded function $\tau_0(u)$ such that uniformly for all $u \in (0, 1)$,

$$\tau_0(u) + o(1) = \frac{\phi(ru)}{\phi(r)}$$

as $r \downarrow 0$ and $1 \leq j \leq 2 + \delta$ with $\delta > 0$, $\int_0^1 (K^j(u))' \tau_0(u) du < \infty$.

(H2) (i) There exist $\beta > 0$ and $\eta_1 > 0$, such that for all $x_1, x_2 \in N_x$, a neighborhood of x , we have

$$|\mathbb{G}(C \mid x_1) - \mathbb{G}(C \mid x_2)| \leq \eta_1 d_{\mathcal{E}}^\beta(x_1, x_2).$$

(H3) There exist $m \geq 2$ and $\eta_2 > 0$, such that, we have, almost surely

$$\mathbb{E}(|Y|^m | X) \leq \eta_2 < \infty;$$

- (i) The conditional mean of $\mathbb{1}_{\{Y_i \in C\}}$ given the σ -field \mathfrak{G}_{i-1} depends only on X_i , i.e., for any $i \geq 1$, $\mathbb{E}(\mathbb{1}_{\{Y_i \in C\}} \mid \mathfrak{G}_{i-1}) = \mathbb{G}(X_i)$ almost surely.
- (ii) The conditional mean of $\mathbb{1}_{\{Y_i \in C\}}$ given the σ -field \mathfrak{G}_{i-1} depends only on X_i , i.e., for any $i \geq 1$, $\mathbb{E}((\mathbb{1}_{\{Y_i \in C\}} - \mathbb{G}(X_i))^2 \mid \mathfrak{G}_{i-1}) = W_2(X_i)$ almost surely. Moreover, the function W_2 is continuous in a neighborhood of x , that is,

$$\sup_{\{u: d(x,u) \leq h\}} |W_2(u) - W_2(x)| = o(1) \quad \text{as } h \rightarrow 0;$$

(H4) For all $(y_1, y_2) \in \mathbb{R}^{2d}$ and constants $b_3 > 0, \eta_4 > 0$, we have for the conditional density $f(\cdot)$ of Y given $X = x$ the following

$$|f(y_1) - f(y_2)| \leq \eta_4 \|y_1 - y_2\|^{b_3};$$

- (i) $F(u; x) = \phi(u)f_1(x)$ as $u \rightarrow 0$, where $\phi(0) = 0$ and $\phi(u)$ is absolutely continuous in a neighborhood of the origin,

(H5) The kernel function $K(\cdot)$ is supported within $(0, 1)$ and has a continuous first derivative on $(0, 1)$ and satisfied the condition $K'(t) < 0 \quad \forall t \in (0, 1)$. Moreover,

$$\left| \int_0^1 (K^j)'(u) du \right| < \infty, \quad \text{for } j = 1, 2.$$

(H6) Assume the class of sets \mathcal{C} satisfies the condition (3);

(H7) The smoothing parameter (h_n) satisfies:

- (i) $\frac{\log n}{n \min(a_n, \phi(h_n))} \rightarrow 0$,
- (ii) Let $h_n \rightarrow 0$ and $n\phi(h_n) \rightarrow \infty$ as $n \rightarrow \infty$.

2.2. Comments on the assumptions

The condition **(H1)** plays an important role in the ergodic and functional context of this paper condition used here share some similarities with that used in [40]. Conditions **(H2)**(i) are classical in the nonparametric regression estimation. **(H3)** is necessary to establish consistency. The condition **(H4)** on the density $f(\cdot)$ is a classical Lipschitz-type nonparametric functional model. **(H5)** The conditions on the kernel are not very restrictive. **(H7)** rules out too large or too small bandwidths without the consistency that could not be obtained. using also the condition (A4) part(ii) of [40] stand as regularity conditions that are of usual nature it is important condition to show the consistency that is omitted here.

3. Main results

Below, we write $Z \stackrel{\mathcal{D}}{=} \mathcal{N}(\mu, \sigma^2)$ whenever the random variable Z follows a normal law with expectation μ and variance σ^2 , $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution and $\xrightarrow{\mathbb{P}}$ the convergence in probability.

Theorem 1. *[Uniform Consistency] Suppose that the hypotheses **(H1)**-**(H7)** hold. Let \mathcal{C} be a class of measurable sets for which*

$$\mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}(\cdot | x)) < \infty,$$

for any $\varepsilon > 0$. Suppose further that $\forall C \in \mathcal{C}$

$$|\mathbb{G}(C, y)f(y) - \mathbb{G}(C, x)f(x)| \longrightarrow 0, \text{ as } y \rightarrow x.$$

If $n\phi(h_n) \rightarrow \infty$ and $h_n \rightarrow 0$ as $n \rightarrow \infty$, then

$$\sup_{C \in \mathcal{C}} |\mathbb{G}_n(C, x) - \mathbb{E}(\mathbb{G}_n(C, x))| \xrightarrow{\mathbb{P}} 0.$$

Remark that, the proof of Theorem 1 is a direct consequence of the decomposition

$$\mathbb{G}_n(C, x) - \mathbb{E}(\mathbb{G}_n(C, x)) = \frac{1}{\mathbb{E}(\widehat{f}_n(x))} \left[\widehat{F}_n(C, x) - \mathbb{E}(\widehat{F}_n(C, x)) \right]$$

$$-\frac{\mathbb{G}_n(C, x)}{\mathbb{E}(\widehat{f}_n(x))} \left[\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x)) \right],$$

where

$$\begin{aligned} \widehat{F}_n(C, x) &= \frac{1}{n\phi(h_n)} \sum_{i=1}^n \mathbb{1}_{\{Y_i \in C\}} K \left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n} \right), \\ \widehat{f}_n(x) &= \frac{1}{n\phi(h_n)} \sum_{i=1}^n K \left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n} \right). \end{aligned}$$

Putting $\Delta_i(x) = K \left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n} \right)$. We have

$$\begin{aligned} \widehat{F}_n(C, x) &= \frac{1}{n\phi(h_n)} \sum_{i=1}^n \mathbb{1}_{\{Y_i \in C\}} \Delta_i(x), \\ \widehat{f}_n(x) &= \frac{1}{n\phi(h_n)} \sum_{i=1}^n \Delta_i(x). \end{aligned}$$

From now for $x \in \mathcal{E}$, set

$$\mathbb{E}(\widehat{F}_n(C, x)) = \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^n \mathbb{E}(\mathbb{1}_{\{Y_i \in C\}} \Delta_i(x) \mid \mathcal{F}_{i-1}),$$

and

$$\mathbb{E}(\widehat{f}_n(x)) = \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^n \mathbb{E}(\Delta_i(x) \mid \mathcal{F}_{i-1}),$$

where $\mathbb{E}(X \mid \mathcal{F})$ is the conditional expectation of the random variables X given the σ -field \mathcal{F} . Lemmas 1 and 2 are important steps towards Theorem 1, for which the proofs are given in the Appendix.

Lemma 1. *Suppose that the hypotheses (H1)-(H7) hold and for every fixed $C \in \mathcal{C}$ as $n \rightarrow \infty$ we have :*

$$\sup_{C \in \mathcal{C}} \left| \widehat{F}_n(C, x) - \mathbb{E} \left(\widehat{F}_n(C, x) \right) \right| = o_{\mathbb{P}}(1).$$

Lemma 2. *Suppose that the hypotheses (H1)-(H7) hold and for every fixed*

$N_{\mathcal{E}}$ neighborhood of x in the functional space \mathcal{E} as $n \rightarrow \infty$, we have

$$\sup_{x \in N_{\mathcal{E}}} \left| \widehat{f}_n(x) - \mathbb{E} \left(\widehat{f}_n(x) \right) \right| = o_{\mathbb{P}}(1).$$

To establish the asymptotic normality define the “bias” term by

$$\begin{aligned} B_n(x) &= \frac{\mathbb{E} \left(\widehat{f}_n(x) \right) - \mathbb{G}_n(C, x) \mathbb{E} \left(\widehat{F}_n(C, x) \right)}{\mathbb{E} \left(\widehat{F}_n(C, x) \right)} \\ &= M_n(x) - \mathbb{G}_n(C, x), \end{aligned} \tag{5}$$

where

$$M_n(x) = \frac{\mathbb{E} \left(\widehat{f}_n(x) \right)}{\mathbb{E} \left(\widehat{F}_n(C, x) \right)}.$$

By stationarity of order one of the (X_i) ’s, we have

$$\mathbb{E}(\widehat{f}_n(x)) = 1. \tag{6}$$

The following result gives the weak convergence. Keep in mind that $f_1(x)$ is given in **(H1)**.

Theorem 2 (Asymptotic normality). *Let **(H1)**-**(H7)** hold. Then as $n \rightarrow \infty$, for $m \geq 1$ and $C_1, \dots, C_m \in \mathcal{C}$,*

$$\{\tilde{\nu}_n(C_i, x)_{i=1, \dots, m}\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma),$$

where $\Sigma = \sigma_{ij}(x), i, j = 1, \dots, m$ and

$$\sigma_{ij}(x) = \frac{\mathfrak{C}_2}{\mathfrak{C}_1^2 f_1(x)} W_2(x),$$

whenever $f_1(x) > 0$ and

$$\mathfrak{C}_1 = k(1) - \int_0^1 K'(u) \tau_0(u) du, \quad \mathfrak{C}_2 = K^2(1) - \int_0^1 (K^2)'(u) \tau_0(u) du.$$

To establish the density of the process, we need to introduce the following function which provides the information on the asymptotic behaviour of

the modulus of continuity

$$\Lambda_\gamma(\sigma^2, n) = \begin{cases} \sqrt{\sigma^2 \log \frac{1}{\sigma^2}}, & \text{if } \gamma = 0; \\ \max \left((\sigma^2)^{(1-\gamma)/2}, n\phi(h_n)^{(3\gamma-1)/(2(3\gamma+1))} \right), & \text{if } \gamma > 0. \end{cases}$$

Theorem 3. *Suppose that (H1)-(H7) hold. For each $\sigma^2 > 0$, let $\mathcal{C}_\sigma \subset \mathcal{C}$ be a class of measurable sets with*

$$\sup_{C \in \mathcal{C}_\sigma} \mathbb{G}(C, x) \leq \sigma^2 \leq 1,$$

and suppose that \mathcal{C} fulfils (3) with $\gamma \geq 0$. Further, we assume that $\phi(h_n) \rightarrow 0$ and $n\phi(h_n) \rightarrow +\infty$ as $n \rightarrow +\infty$, such that

$$n\phi(h_n) \leq (\Lambda_\gamma(\sigma^2, n))^2,$$

and as $n \rightarrow +\infty$, we have

$$\frac{n\phi \left(\sigma^2 \log \left(\frac{1}{\sigma^2} \right) \right)^{1+\gamma}}{\log(n)} \rightarrow \infty.$$

Further we assume that $\sigma^2 \geq h^2$. For $\gamma > 0$ and $d = 1, 2$, the later has to be replaced by $\sigma^2 \geq \phi(h_n) \log \left(\frac{1}{\phi(h_n)} \right)$, then under conditions of Theorem 2 we have the process:

$$\{\tilde{\nu}_n(C, x) : C \in \mathcal{C}\},$$

converges in law to a Gaussian process $\{\tilde{\nu}(C, x) : C \in \mathcal{C}\}$, that admits a version with uniformly bounded and uniformly continuous paths with respect to $\|\cdot\|_2$ -norm with covariance $\sigma_{ij}(x)$ given in Theorem 2.

Remark 1. Central limit theorems are usually used to establish confidence intervals for the target to be estimated. In the context of non-parametric estimation the asymptotic variance $\Sigma(x) := \sigma_{i,j}(x)$ in the central limit depends on certain functions only approximate confidence intervals can be obtained in practice, even when $\Sigma(x)$ functionally specified. Observing now in (2) that the limiting variance contains the unknown function f_1 and that the normalization depends on the function $\phi(\cdot)$ which is not identifiable explicitly. Moreover, we have to estimate the quantities W_2 and τ_0 the corollary

below is a slight modification of (2) allows to have usable form of our results in practice as usually the conditional variance $W_2(x)$ is estimated by

$$\begin{aligned} W_{2,n} &= \frac{\sum_{i=1}^n (\mathbb{1}_{\{Y_i \in C\}} - \mathbb{G}_n(x))^2 K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h}\right)}{\sum_{i=1}^n K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h}\right)} \\ &= \frac{\sum_{i=1}^n (\mathbb{1}_{\{Y_i \in C\}} - \mathbb{G}_n(x))^2 K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h}\right)}{\sum_{i=1}^n K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h}\right)} - (\mathbb{G}_n(x))^2 \\ &= \hat{g}_n(x) - (\mathbb{G}_n(x))^2. \end{aligned}$$

Let us introduce the following estimate

$$\mathcal{F}_{x,n}(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{d(x, X_i) \leq t\}}.$$

Making use the decomposition of $\tau_0(\cdot)$ in (H1)(i) one may estimate $\tau_0(\cdot)$ by

$$\tau_n(t) = \frac{\mathcal{F}_{x,n}(th)}{\mathcal{F}_{x,n}(h)}.$$

Subsequently, for a given kernel $K(\cdot)$ and the quantities \mathfrak{C}_1 and \mathfrak{C}_2 can be estimated as follows

$$\mathfrak{C}_{1,n} = K(1) - \int_0^1 K'(s)\tau_n(s)ds, \quad \mathfrak{C}_{2,n} = K^2(1) - \int_0^1 (K^2)'(s)\tau_n(s)ds.$$

Introduce now some further conditions needed to state

- (H8) (i) The conditional mean of $\mathbb{1}_{\{Y_i^2 \in C\}}$ given the σ -field \mathfrak{G}_{i-1} depends only on X_i , i.e., there exist a function g such that for any $i \geq 1$, $\mathbb{E}(\mathbb{1}_{\{Y_i^2 \in C\}} \mid \mathfrak{G}_{i-1}) = g(X_i)$ almost surely,
- (ii) The conditional variance of $\mathbb{1}_{\{Y_i^2 \in C\}}$ given \mathfrak{G}_{i-1} depends only on X_i i.e., for any $i \geq 1$ $\mathbb{E}\left(\left(\mathbb{1}_{\{Y_i^2 \in C\}}\right)^2 \mid \mathfrak{G}_{i-1}\right) = U(X_i)$ almost surely, for some function U . Moreover, the function U is continuous in a

neighborhood of x , that is

$$\sup_{u:d(x,u)\leq h} |U(u) - U(x)| = o(1).$$

Corollary 3.1. *Assume that conditions (H1)-(H8) hold true K' and $(K^2)'$ are integrable functions and $n\mathcal{F}_x(h) \rightarrow \infty$ and $h^\beta(n\mathcal{F}_x(h))^{1/2} \rightarrow 0$ as $n \rightarrow \infty$. Then, for any $x \in \mathcal{E}$ such that $f_1(x) > 0$, we have*

$$\frac{\mathfrak{C}_{1,n}}{\sqrt{\mathfrak{C}_{2,n}}} \sqrt{\frac{n\mathcal{F}_{x,n}(h_n)}{W_{2,n}(x)}} (\mathbb{G}_n(C, x) - \mathbb{G}(C, x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Using Corollary (3.1) the asymptotic $100(1 - \alpha)$ confidence band given by

$$\left[\mathbb{G}_n(C, x) - c_\alpha \frac{\mathfrak{C}_{1,n}}{\sqrt{\mathfrak{C}_{2,n}}} \sqrt{\frac{W_{2,n}(x)}{n\mathcal{F}_{x,n}(h)}}, \mathbb{G}_n(C, x) + c_\alpha \frac{\mathfrak{C}_{1,n}}{\sqrt{\mathfrak{C}_{2,n}}} \sqrt{\frac{W_{2,n}(x)}{n\mathcal{F}_{x,n}(h)}} \right].$$

3.1. The bandwidth selection criterion

Many methods have been established and developed to construct, in asymptotically optimal ways, bandwidth selection rules for nonparametric kernel estimators especially for Nadaraya-Watson regression estimator we quote among them [32], [34], [47], [22], [13] and [14]. This parameter has to be selected suitably, either in the standard finite dimensional case, or in the infinite dimensional framework for insuring good practical performances. Let us define the leave-out- (X_i, Y_i) estimator for regression function

$$\mathbb{G}_{n,j}(C, x) = \frac{\sum_{i=1, i \neq j}^n \mathbb{1}_{\{Y_i \in C\}} K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}. \tag{7}$$

In order to minimize the quadratic loss function, we introduce the following

criterion, we have for some (known) non-negative weight function $\mathcal{W}(\cdot)$:

$$CV(C, h) := \frac{1}{n} \sum_{j=1}^n \left(\mathbb{1}_{\{Y_j \in C\}} - \mathbb{G}_{n,j}(C, X_j) \right)^2 \mathcal{W}(X_j). \tag{8}$$

Following the ideas developed by [47], a natural way for choosing the bandwidth is to minimize the precedent criterion, so let's choose $\hat{h}_n \in [a_n, b_n]$ minimizing among $h \in [a_n, b_n]$:

$$\sup_{C \in \mathcal{C}} CV(\Psi, h).$$

The main interest of our results is the possibility to derive the asymptotic properties of our estimate even if the bandwidth parameter is a random variable, like in the last equation. One can replace (8) by

$$CV(C, h_n) := \frac{1}{n} \sum_{j=1}^n \left(\mathbb{1}_{\{Y_j \in C\}} - \mathbb{G}_{n,j}(C, X_j) \right)^2 \widehat{\mathcal{W}}(X_j, x). \tag{9}$$

In practice, one takes, for $j = 1, \dots, n$, the uniform global weights $\mathcal{W}(X_j) = 1$, and the local weights

$$\widehat{\mathcal{W}}(X_j, x) = \begin{cases} 1 & \text{if } d(X_j, x) \leq h_n, \\ 0 & \text{otherwise.} \end{cases}$$

For sake of brevity, we have just considered the most popular method, that is, the cross-validated selected bandwidth. This may be extended to any other bandwidth selector such the bandwidth based on Bayesian ideas [51].

4. Testing the Independence

We consider a sample of random elements $(X_1, Y_{1,1}, Y_{1,2}), \dots, (X_n, Y_{n,1}, Y_{n,2})$ copies of (X, Y_1, Y_2) that takes its value in a space $\mathcal{E} \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and define, for $(C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2$,

$$\mathbb{G}_n(C_1 \times C_2, x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{Y_{i,1} \in C_1\}} \mathbb{1}_{\{Y_{i,2} \in C_2\}} K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}, \tag{10}$$

$$\mathbb{G}_{n,1}(C_1, x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{Y_{i,1} \in C_1\}} K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}, \tag{11}$$

$$\mathbb{G}_{n,2}(C_2, x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{Y_{i,2} \in C_2\}} K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{d_{\mathcal{E}}(x, X_i)}{h_n}\right)}. \tag{12}$$

We will investigate the following processes, for $(C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2$,

$$\widehat{\nu}_n(C_1, C_2, x) = \sqrt{n\phi(h_n)} (\mathbb{G}_n(C_1 \times C_2, x) - \mathbb{E}(\mathbb{G}_n(C_1, x))\mathbb{E}(\mathbb{G}_n(C_2, x))), \tag{13}$$

$$\check{\nu}_n(C_1, C_2, x) = \sqrt{n\phi(h_n)} (\mathbb{G}_n(C_1 \times C_2, x) - \mathbb{G}_{n,1}(C_1, x)\mathbb{G}_{n,2}(C_2, x)). \tag{14}$$

Notice that we have

$$\begin{aligned} \check{\nu}_n(C_1, C_2, x) &= \sqrt{n\phi(h_n)} (\mathbb{G}_n(C_1 \times C_2, x) - \mathbb{E}(\mathbb{G}_n(C_1, x))\mathbb{E}(\mathbb{G}_n(C_2, x))) \\ &\quad + \sqrt{n\phi(h_n)}\mathbb{E}(\mathbb{G}_n(C_2, x)) (\mathbb{G}_n(C_1, x) - \mathbb{E}(\mathbb{G}_n(C_1, x))) \\ &\quad - \sqrt{n\phi(h_n)}(\mathbb{G}_n(C_1, x)) (\mathbb{G}_n(C_2, x) - \mathbb{E}(\mathbb{G}_n(C_2, x))). \end{aligned}$$

Hence we have

$$\begin{aligned} \check{\nu}_n(C_1, C_2, x) &\stackrel{d}{=} \sqrt{n\phi(h_n)} (\mathbb{G}_n(C_1 \times C_2, x) - \mathbb{E}(\mathbb{G}_n(C_1, x))\mathbb{E}(\mathbb{G}_n(C_2, x))) \\ &\quad + \sqrt{n\phi(h_n)}\mathbb{E}(\mathbb{G}_n(C_2, x)) (\mathbb{G}_n(C_1, x) - \mathbb{E}(\mathbb{G}_n(C_1, x))) \\ &\quad - \sqrt{n\phi(h_n)}\mathbb{E}(\mathbb{G}_n(C_1, x)) (\mathbb{G}_n(C_2, x) - \mathbb{E}(\mathbb{G}_n(C_2, x))) \\ &= \widehat{\nu}_n(C_1, C_2, x) + \mathbb{E}(\mathbb{G}_n(C_2, x))\widetilde{\nu}_n(C_1, x) - \mathbb{E}(\mathbb{G}_n(C_1, x)) \\ &\quad \times \widetilde{\nu}_n(C_2, x). \end{aligned} \tag{15}$$

Let $\{\widehat{\nu}(C_1, C_2, x) : (C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2\}$ be a Gaussian process. Let us introduce the following limiting process, for $(C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2$,

$$\check{\nu}(C_1, C_2, x) = \widehat{\nu}(C_1, C_2, x) + \mathbb{G}(C_2, x)\widetilde{\nu}(C_1, x) - \mathbb{G}(C_1, x)\widetilde{\nu}(C_2, x).$$

We would test the following null hypothesis

$$\mathcal{H}_0 : Y_1 \text{ and } Y_2 \text{ are conditionally independent given } X = x.$$

Against the alternative

$$\mathcal{H}_1 : Y_1 \text{ and } Y_2 \text{ are conditionally dependent.}$$

Statistics of independence those can be used are

$$S_{1,n} = \sup_{(C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2} |\widehat{\nu}_n(C_1, C_2, x)|, \tag{16}$$

$$S_{2,n} = \sup_{(C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2} |\check{\nu}_n(C_1, C_2, x)|. \tag{17}$$

A combination of Theorem 3 with continuous mapping theorem we obtain the following result.

Theorem 4. *We have under condition of Theorem 3, as $n \rightarrow \infty$,*

$$S_{1,n} \rightarrow \sup_{(C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2} |\widehat{\nu}(C_1, C_2, x)|, \tag{18}$$

$$S_{2,n} \rightarrow \sup_{(C_1, C_2) \in \mathcal{C}_1 \times \mathcal{C}_2} |\check{\nu}(C_1, C_2, x)|. \tag{19}$$

5. Concluding Remarks

In the present work, we have established the invariance principle for the conditional set-indexed empirical process formed by ergodic functional data. Our results are obtained under assumptions on the richness of the index class \mathcal{C} of sets in terms of metric entropy with bracketing in the framework of ergodic variables. This paper extends the dependence setting to the cases not covered by the usual mixing structures because ideas to general dependence structure is a rather underdeveloped field. Note that the ergodic framework avoids the widely used variants to measure the dependency and our work would go well beyond the scope of the empirical process literature, recall that the theory of empirical process is useful in many applications and an application Bahadur presentation we will derive.

6. Appendix

This section is devoted to the proof of our results. The aforementioned notation is also used in what follows.

Proof of Lemma 1. Use finite metric entropy with inclusion, fix $\epsilon > 0$ for $C \in \mathcal{C}$. Let C_*, C^* be a bracket for C , i.e., $C_* \subset C \subset C^*$, such that

$$\mathbb{G}(C_* \Delta C^* \mid x) < \epsilon.$$

Since for $A \subset B$ we have $\mathbb{G}_n(A, x) \leq \mathbb{G}_n(B, x)$ and $\mathbb{G}(A \mid x) \leq \mathbb{G}(B \mid x)$, it follows:

$$\begin{aligned} & \sup_{C \in \mathcal{C}} [\mathbb{G}_n(C, x) - \mathbb{E}(\mathbb{G}_n(C, x))] \\ & \leq \sup_{C \in \mathcal{C}} [\mathbb{G}_n(C^*, x) - \mathbb{E}(\mathbb{G}_n(C_*, x))] \\ & \leq \sup_{C \in \mathcal{C}} [\mathbb{G}_n(C^*, x) - \mathbb{E}(\mathbb{G}_n(C^*, x))] + \sup_{C \in \mathcal{C}} [\mathbb{E}(\mathbb{G}_n(C^*, x)) - \mathbb{E}(\mathbb{G}_n(C_*, x))] \\ & \leq \sup_{C \in \mathcal{C}} [\mathbb{G}_n(C^*, x) - \mathbb{E}(\mathbb{G}_n(C^*, x))] + \sup_{C \in \mathcal{C}} \mathbb{G}(C_* \Delta C^* \mid x) \\ & \leq \sup_{C \in \mathcal{C}} [\mathbb{G}_n(C^*, x) - \mathbb{E}(\mathbb{G}_n(C^*, x))] + \epsilon. \end{aligned} \tag{20}$$

An analogous lower bound holds with C^* replaced by C_* . Since the first term in the last line is a supremum over finitely sets (for fixed $\epsilon > 0$) it follows pointwise consistency of $\mathbb{G}_n(\cdot, \cdot)$ that the term is $o_{\mathbb{P}}(1)$ and hence we obtain the desired result. \square

Lemma 3. Assume that condition **(H1(i))**- **(H1(ii))**-**(H1(iv))**-**(H6)** hold true for any real numbers $1 \leq j \leq 2 + \delta$ and $1 \leq k \leq 2 + \delta$ with $\delta > 0$ as $n \rightarrow \infty$ we have :

- (i) $\frac{1}{\phi(h)} \mathbb{E}(\Delta_i^j(x) \mid \mathfrak{F}_{i-1}) = M_j f_{i,1}(x) + \mathcal{O}_{a.s} \left(\frac{g_{i,x}(h)}{\phi(h)} \right).$
- (ii) $\frac{1}{\phi(h)} \mathbb{E}(\Delta_1^j(x)) = M_j f_1(x) + o(1).$
- (iii) $\frac{1}{\phi^k(h)} (\mathbb{E}(\Delta_1(x)))^k = M_1^k f_1^k(x) + o(1).$

The reader is referred to [40] for the proof of Lemma 3. \square

Proof of Lemma 2. We shall proof that

$$\mathbb{P} \left(\left| \widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n) \right| > \epsilon \right) \rightarrow 0.$$

Observe the condition in (6) and we use the same proof in [40] look that $\widehat{f}_n(x) - 1 = R_{1,n}(x) + R_{2,n}(x)$ where

$$\begin{aligned}
 R_{1,n}(x) &= \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{t=1}^n (\Delta_t(x) - \mathbb{E}(\Delta_t(x) \mid \mathfrak{F}_{i-1})) \\
 R_{2,n}(x) &= \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{t=1}^n (\mathbb{E}[\Delta_t(x) \mid \mathfrak{F}_{i-1}] - \mathbb{E}(\Delta_1(x))) \\
 &= \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{t=1}^n \mathbb{E}[\Delta_t(x) \mid \mathfrak{F}_{i-1}] - 1.
 \end{aligned}$$

Combining 3 with hypothesis (H1)-(ii) and (H1)-(iii) it easy seen that $R_{2,n}(x) = o_{a.s}(1)$ as $n \rightarrow \infty$. For the first term observe that $R_{1,n}(x) = \sum_{t=1}^n L_{ni}(x)$, where $L_{ni}(x)$ is a triangular array of martingale differences with respect to the σ -field \mathfrak{F}_{i-1} combining Burkholder [33] and Jensen inequalities, we obtain for any $\epsilon > 0$ that exists a constant C_0 such that

$$\mathbb{P}(|R_{1,n}(x)| > \epsilon) \leq C_0 \frac{\mathbb{E}(\Delta_1^2(x))}{\epsilon^2 n (\mathbb{E}(\Delta_1(x)))^2} = \mathcal{O}\left(\frac{1}{\epsilon^2 n \phi(h)} + o(1)\right),$$

where the last equality results form (3). Since $n\phi(h) \rightarrow \infty$ as $n \rightarrow \infty$ we conclude then that $R_{1,n}(x) = o_{\mathbb{P}}(1)$. Thus the proof is complete. \square

Proof of Theorem 2. We will use similar arguments to those used in the paper by [40] to prove the asymptotic normality of the process we shall use the following notation, recall the decomposition:

$$\begin{aligned}
 \mathbb{G}_n(C, x) - \mathbb{E}(\mathbb{G}_n(C, x)) &= \frac{1}{\mathbb{E}(\widehat{f}_n(x))} \left[\widehat{F}_n(C, x) - \mathbb{E}(\widehat{F}_n(C, x)) \right] \\
 &\quad - \frac{\mathbb{G}_n(C, x)}{\mathbb{E}(\widehat{f}_n(x))} \left[\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x)) \right] \\
 &= \frac{Q_n(x)}{\mathbb{E}(\widehat{f}_n(x))},
 \end{aligned}$$

where

$$Q_n(x) = \left[\widehat{F}_n(C, x) - \mathbb{E}(\widehat{F}_n(C, x)) \right] - \mathbb{G}_n(C, x) \left[\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x)) \right].$$

Lemma 4. *Assume that the hypotheses (H1)-(H7) are satisfied, then we have for any $x \in \mathcal{E}$ such that $f_1(x) > 0$, we have :*

$$\sqrt{n\phi(h_n)}Q_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \text{as } n \rightarrow \infty.$$

Proof of Lemma 4. Let us introduce some notation. Set

$$\eta_{ni} = \left(\frac{\phi(h)}{n}\right)^{1/2} (\mathbb{1}_{\{Y_i \in C\}} - \mathbb{G}(x)) \frac{\Delta_i(x)}{\mathbb{E}(\Delta_1(x))}, \tag{21}$$

and define $\xi_{ni} = \eta_{ni} - \mathbb{E}(\eta_{ni} \mid \mathfrak{F}_{i-1})$. It is easily seen that

$$(n\phi(h))^{1/2}Q_n(x) = \sum_{t=1}^n \xi_{ni}, \tag{22}$$

where for any fixed $x \in \mathcal{E}$ the summands (22) form a triangular array of stationary martingale differences with respect to the σ -field \mathfrak{F}_{i-1} . This allows us to apply the central limit theorem for discrete-time arrays of real-valued martingales (see, [31] page 23) to establish the asymptotic normality of $Q_n(x)$. This can be done if we establish the following statements:

$$(a) \quad \sum_{t=1}^n \mathbb{E}(\xi_{ni}^2 \mid \mathfrak{F}_{i-1}) \longrightarrow \sigma^2(x),$$

and

$$(b) \quad n\mathbb{E}(\xi_{ni}^2 \mathbb{1}_{|\eta_{ni}| > \epsilon}) = o(1),$$

holds for any $\epsilon > 0$ (Lindeberg condition).

Proof of Part (a). Observe first that

$$\left| \sum_{t=1}^n \mathbb{E}(\eta_{ni}^2 \mid \mathfrak{F}_{i-1}) - \sum_{t=1}^n \mathbb{E}(\xi_{ni}^2 \mid \mathfrak{F}_{i-1}) \right| \leq \sum_{t=1}^n (\mathbb{E}(\eta_{ni} \mid \mathfrak{F}_{i-1}))^2.$$

Making use of the condition (H2) and Lemma 3, one has

$$\begin{aligned} \mathbb{E}(\eta_{ni} \mid \mathfrak{F}_{i-1}) &= \frac{1}{\mathbb{E}(\Delta_i)} \left(\frac{\phi(h)}{n}\right)^{1/2} |\mathbb{E}((\mathbb{G}(X_i) - \mathbb{G}(x))\Delta_i(x) \mid \mathfrak{F}_{i-1})| \\ &\leq \frac{1}{\mathbb{E}(\Delta_i)} \left(\frac{\phi(h)}{n}\right)^{1/2} \sup_{u \in B(x,h)} |\mathbb{G}(X_i) - \mathbb{G}(x)| \mathbb{E}(\Delta_i(x) \mid \mathfrak{F}_{i-1}) \end{aligned}$$

$$= \mathcal{O}(h^\beta) \left(\frac{\phi(h)}{n} \right)^{1/2} \left(\frac{f_{i,1}(x)}{f_1(x)} + \mathcal{O}_{a.s} \left(\frac{g_{i,x}(h)}{\phi(h)} \right) \right). \tag{23}$$

Thus, by (H1)(ii)(iii), we have

$$\begin{aligned} \sum_{t=1}^n (\mathbb{E}(\eta_{mi} \mid \mathfrak{F}_{i-1}))^2 &= \mathcal{O}(h^{2\beta}) \left(\frac{\phi(h)}{n} \right) \sum_{t=1}^n \left(\frac{f_{i,1}(x)}{f_1(x)} + \mathcal{O}_{a.s} \left(\frac{g_{i,x}(h)}{\phi(h)} \right) \right)^2 \\ &= \mathcal{O}(h^{2\beta} \phi(h)) \left(\frac{1}{f_1^2(x)} \frac{1}{n} \sum_{t=1}^n f_{i,1}^2(x) + o_{a.s}(1) \right) \\ &= \mathcal{O}_{a.s}(\phi(h)h^{2\beta}). \end{aligned} \tag{24}$$

The statement (a) follows then if we show that

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n \mathbb{E}(\eta_{mi}^2 \mid \mathfrak{F}_{i-1}) = \sigma^2. \tag{25}$$

To prove (25), observe that

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{t=1}^n \mathbb{E}(\eta_{mi}^2 \mid \mathfrak{F}_{i-1}) &= \frac{\phi(h)}{n(\mathbb{E}(\Delta_1(x)))^2} \sum_{t=1}^n \mathbb{E}[(\mathbb{1}_{\{Y_i \in C\}} - \mathbb{G}(x))^2 \Delta_i^2(x) \mid \mathfrak{F}_{i-1}] \\ &= J_{1n} + J_{2n}, \end{aligned}$$

where

$$\begin{aligned} J_{1n} &= \frac{\phi(h)}{n(\mathbb{E}(\Delta_1(x)))^2} \sum_{t=1}^n \mathbb{E} \left[\Delta_i^2(x) \mathbb{E}(\mathbb{1}_{\{Y_i \in C\}} - \mathbb{G}(x))^2 \mid \mathfrak{G}_{i-1} \mid \mathfrak{F}_{i-1} \right] \\ &= \frac{\phi(h)}{n(\mathbb{E}(\Delta_1(x)))^2} \sum_{t=1}^n \mathbb{E} [W_2(X_i) \Delta_i^2(x) \mid \mathfrak{F}_{i-1}], \end{aligned} \tag{26}$$

and

$$J_{2n} = \frac{\phi(h)}{n(\mathbb{E}(\Delta_1(x)))^2} \sum_{t=1}^n \mathbb{E} [(\mathbb{G}(X_i) - \mathbb{G}(X))^2 \Delta_i^2(x) \mid \mathfrak{F}_{i-1}].$$

We give now an upper bound for

$$\mathbb{E} [W_2(X_i) \Delta_i^2(x) \mid \mathfrak{F}_{i-1}].$$

Towards this end, we split it up into

$$I_{n1} + I_{n2},$$

with

$$I_{n1} = W_2(x)\mathbb{E}(\Delta_i^2(x) \mid \mathfrak{F}_{i-1})$$

and

$$I_{n2} = \mathbb{E} [(W_2(X_i) - W_2(x))\Delta_i^2(x) \mid \mathfrak{F}_{i-1}]$$

Using the condition (A4) part(ii) of [40] , one can write

$$|I_{n2}| \leq \sup_{u:d(x;u)\leq h} |W_2(u)-W_2(x)|\mathbb{E} [\Delta_i^2(x) \mid \mathfrak{F}_{i-1}] = \mathbb{E} [\Delta_i^2(x) \mid \mathfrak{F}_{i-1}] \times o(1).$$

Thus, in view of 3 part (i), we have

$$\begin{aligned} \mathbb{E} [W_2(X_i)\Delta_i^2(x) \mid \mathfrak{F}_{i-1}] &= (o(1) + W_2(x))\mathbb{E} (\Delta_i^2(x) \mid \mathfrak{F}_{i-1}) \\ &= (o(1)+W_2(x))(M_2\phi(h)f_{i,1}(x)+O_{a.s}(g_{i,x}(h))). \end{aligned} \tag{27}$$

Combining again 3 and conditions (H1)(ii)(iii), it is easily seen that $\lim_{n \rightarrow \infty} J_{1n} = \frac{M_2}{M_1^2} \frac{W_2(x)}{f_1(x)}$ almost surely, whenever $f_1(x) > 0$. Consider now the term J_{2n} . Making use of conditions (H1)(ii)(iii) and (H2)-(i) and Lemma 3, one can write

$$\begin{aligned} |J_{n2}| &= \mathcal{O}(h^{2\beta}) \frac{\phi(h)}{n(\mathbb{E}(\Delta_1(x)))^2} \sum_{t=1}^n \mathbb{E} (\Delta_i^2(x) \mid \mathfrak{F}_{i-1}) \\ &= \mathcal{O}(h^{2\beta}) \left(\frac{M_2}{M_1^2} \frac{1}{f_1(x)} + o_{a.s}(1) \right) \rightarrow 0 \text{ almost surely as } n \rightarrow \infty. \end{aligned} \tag{28}$$

Therefore,

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n \mathbb{E} (\Delta_i^2(x) \mid \mathfrak{F}_{i-1}) = \lim_{n \rightarrow \infty} (J_{n1} + J_{n2}) = \frac{M_2}{M_1^2} \frac{W_2(x)}{f_1(x)} =: \sigma^2(x)$$

almost surely whenever $f_1(x) > 0$, this completes the proof of Part (a). \square

Proof of Part (b). The Lindeberg condition results from Corollary 9.5.2 in ([21]) which implies that

$$n\mathbb{E}(\xi_{ni}^2 \mathbb{1}(|\xi_{ni}| > \varepsilon)) \leq 4n\mathbb{E}(\eta_{ni}^2 \mathbb{1}(|\eta_{ni}| > \varepsilon/2)).$$

Let $a > 1$ and $b > 1$ such that $\frac{1}{a} + \frac{1}{b} = 1$. Making use of Hlder and Markov

inequalities one can write, for all $\varepsilon > 0$.

$$\mathbb{E}(\eta_{mi}^2 \mathbf{1}(|\eta_{mi}| > \varepsilon/2)) \leq \frac{\mathbb{E}|\eta_{mi}|^{2a}}{(\varepsilon/2)^{2a/b}},$$

taking C_0 a positive constant and $2a = 2 + \delta$ (with δ using the condition (A4) part(ii) of [40]), we obtain

$$\begin{aligned} & 4n\mathbb{E}(\eta_{mi}^2 \mathbf{1}(|\eta_{mi}| > \varepsilon/2)) \\ \leq & C_0 \left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2} \frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}} \mathbb{E}([\mathbf{1}_{\{Y_i \in C\}} - \mathbf{G}(x)|\Delta_i(x)]^{2+\delta}) \\ \leq & C_0 \left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2} \frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}} \mathbb{E}\left(\mathbb{E}\left(|\mathbf{1}_{\{Y_i \in C\}} - \mathbf{G}(x)|^{2+\delta}(\Delta_i(x))^{2+\delta} \mid X_i\right)\right) \\ \leq & C_0 \left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2} \frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}} \mathbb{E}\left((\Delta_i(x))^{2+\delta} \overline{W}_{2+\delta}(X_i)\right) \\ \leq & C_0 \left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2} \frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}} \left(\mathbb{E}\left((\Delta_i(x))^{2+\delta} |\overline{W}_{2+\delta}(X_i) - \overline{W}_{2+\delta}(x)\right)\right. \\ & \left. + |\overline{W}_{2+\delta}(x)| \mathbb{E}\left[(\Delta_i(x))^{2+\delta}\right]\right) \\ \leq & C_0 \left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2} \frac{n\mathbb{E}\left[(\Delta_1(x))^{2+\delta}\right]}{\mathbb{E}(\Delta_1(x))^{2+\delta}} (|\overline{W}_{2+\delta}(x)| + o(1)) \\ \leq & C_0(n\phi(h))^{-\delta/2} \frac{(M_{2+\delta}f_1(x) + o(1))}{(M_1^{2+\delta}f_1^{2+\delta}(x) + o(1))} (|\overline{W}_{2+\delta}(x)| + o(1)) \\ = & \mathcal{O}((n\phi(h))^{-\delta/2}), \end{aligned}$$

where the last equality follows from Lemma 3. This completes the proof of part (b), since $n\phi(h) \rightarrow \infty$ as $n \rightarrow \infty$. Thus the proof is complete. \square

Proof of Theorem 3. Let us recall some facts. Let $f(\cdot) = \mathbf{1}\{\cdot \in C_1\}$ and $g(\cdot) = \mathbf{1}\{\cdot \in C_2\}$. Given random measures μ_n on $(\mathbf{X}, \mathcal{X})$, we define

$$d_{\mu_n}^{(2)}(f, g) := [\mu_n(f - g)^2]^{1/2}.$$

Say that a class of functions \mathcal{F} has uniformly integrable entropy with respect to \mathcal{L}_2 -norm if

$$\int_0^\infty \sup_{\gamma \in M(\mathbf{X}, \mathcal{F})} \left[\ln N\left(\epsilon [\gamma(F^2)]^{1/2}, \mathcal{F}, d_\gamma^{(2)}\right) \right]^{1/2} d\epsilon < \infty,$$

where $d_\gamma^{(2)}(f, g) := [\int_{\mathbf{X}} (f - g)^2 d\gamma]^{1/2}$. When the class \mathcal{F} has uniformly integrable entropy, $(\mathcal{F}, d_\gamma^{(2)})$ is totally bounded for any measure γ . Let κ be an envelope of \mathcal{F} . That is, κ a measurable function mapping \mathcal{F} to $[0, \infty)$ such that

$$\sup_{f \in \mathcal{F}} |f(t)| \leq \kappa(t), \quad \text{for all } t \in \mathbb{R}.$$

Let $M(\mathbb{R}, \kappa)$ be the set of all measures γ on $(\mathbb{R}, \mathcal{F})$ with

$$\gamma(\kappa) := \int_{\mathbb{R}} \kappa^2 d\gamma < \infty, \tag{29}$$

and

$$d_\gamma^{(r)}(f, g) := \left[\int_{\mathbb{R}} (f - g)^r d\gamma \right]^{1/r}.$$

Given random measures μ_n on $(\mathbb{R}, \mathcal{F})$, we define

$$d_{\mu_n}^{(2)}(f, g) := [\mu_n(f - g)^2]^{1/2}.$$

Let us introduce the uniform entropy integral

$$J(\delta, \mathcal{F}, d_\gamma^{(2)}) = \int_0^\delta \sup_{\gamma \in (\mathbb{R}, \mathcal{F})} \left[\log \left(\mathcal{N} \left(\epsilon [\gamma(\kappa^2)]^{1/2}, \mathcal{F}, d_\gamma^{(2)} \right) \right) \right]^{1/2} d\epsilon.$$

We say that \mathcal{F} has uniformly integrable entropy with respect to L_2 -norm if

$$J(\infty, \mathcal{F}, d_\gamma^{(2)}) < \infty. \tag{30}$$

When the class \mathcal{F} has uniformly integrable entropy, $(\mathcal{F}, d_\gamma^{(2)})$ is totally bounded for any measure γ . Let $\{\mathbb{B}(\varphi) : \varphi \in \mathcal{F}\}$ be a Gaussian process whose sample paths are contained in

$$U_b(\mathcal{F}, d_\gamma^{(2)}) := \left\{ f \in \ell^\infty(\mathcal{F}) : f \text{ is uniformly continuous with respect to } d_\gamma^{(2)} \right\}.$$

Let $\mathcal{L}(\bullet)$ denote the law of \bullet . Notice that obtaining a uniform CLT essentially means that we show the following convergence

$$\left\{ \mathcal{L}(A_{n,\varphi}) : \varphi \in \mathcal{F} \right\} \rightarrow \left\{ \mathcal{L}(\mathbb{B}(\varphi)) : \varphi \in \mathcal{F} \right\},$$

where the processes are indexed by \mathcal{F} and considered as random elements of the bounded real-valued functions on \mathcal{F} defined by

$$\ell^\infty(\mathcal{F}) := \left\{ f : \mathcal{F} \rightarrow \mathbb{R} : \|f\|_{\mathcal{F}} := \sup_{\varphi \in \mathcal{F}} |f(\varphi)| < \infty \right\}, \tag{31}$$

which is a Banach space equipped with the sup norm. In the sequel, we use the weak convergence in the sense of [35] that we recall in the following definition. Throughout the paper, \mathbb{E}^* denotes the upper expectation with respect to the outer probability \mathbb{P}^* , we refer to [55, p.6] and [39, §6.2, p.88] for further details and discussion.

Definition 1. A sequence of $\ell^\infty(\mathcal{F})$ -valued random functions $\{T_n : n \geq 1\}$ converges in law to a $\ell^\infty(\mathcal{F})$ -valued Borel measurable random function T whose law concentrates on a separable subset of $\ell^\infty(\mathcal{F})$, denoted $T_n \rightsquigarrow T$, if,

$$\mathbb{E}g(T) = \lim_{n \rightarrow \infty} \mathbb{E}^*g(T_n), \quad \forall g \in C(\ell^\infty(\mathcal{F}), \|\cdot\|_{\mathcal{F}}),$$

where $C(\ell^\infty(\mathcal{F}), \|\cdot\|_{\mathcal{F}})$ is the set of all bounded $\|\cdot\|_{\mathcal{F}}$ -continuous functions from $(\ell^\infty(\mathcal{F}), \|\cdot\|_{\mathcal{F}})$ into \mathbb{R} .

We set

$$\eta_{n;i}(f, x) := \eta_{n;i}(C_1, x) := \left(\frac{\phi(h)}{n} \right)^{1/2} \left(\mathbb{1}_{\{Y_i \in C_1\}} - \mathbb{G}(C, x) \right) \frac{\Delta_i(x)}{\mathbb{E}(\Delta_i(x))}.$$

with $\Delta_i(x) = K(h^{-1}d(x, X_i))$, and define $\eta_{n;i}(g, x)$ in a similar way. Let

$$\xi_{n;i}(f, x) := \eta_{n;i}(f, x) - \mathbb{E}(\eta_{n;i}(f, x) \mid \mathfrak{F}_{i-1}).$$

Let us define

$$\sigma_n^2(f, g) = \sum_{i=1}^n (\xi_{n;i}(f, x) - \xi_{n;i}(g, x))^2.$$

To prove Theorem 3, using Theorem 2 of [4], it suffices to show that, for all constant $L > 0$, as n tends to infinity, that

$$\mathbb{P}^* \left\{ \sup_{f, g \in \mathcal{F}} \frac{\sigma_n^2(f, g)}{(d_{\mu_n}^{(2)}(f, g))^2} > L \right\} \rightarrow 0, \tag{32}$$

which is implied by the following,

$$\mathbb{E}^* \sup_{d^{(2)}(f,g) \leq \delta_n} \sum_{i=1}^n \frac{\mathbb{E}((\xi_{n;i}(f, x) - \xi_{n;i}(g, x))^2 \mid \mathfrak{F}_{i-1})}{(d^{(2)}(f, g))^2} \rightarrow 0, \text{ as } \delta_n \rightarrow 0,$$

where we recall

$$d^{(2)}(f, g) := \left[\int_{\mathbb{R}} (f - g)^2 d\mathbb{P} \right]^{1/2}.$$

In the rest of the proof, denote by $\beta_n(x) = \frac{\sqrt{\phi(h)}}{\mathbb{E}[\Delta_1(x)]}$, and

$$\zeta(f, x) = \zeta(C_1, x) := (\mathbb{1}_{\{Y_i \in C_1\}} - \mathbb{G}(C, x)) \Delta_i(x).$$

Therefore, we have the following

$$\begin{aligned} & \sum_{i=1}^n \frac{\mathbb{E}((\xi_{n;i}(f, x) - \xi_{n;i}(g, x))^2 \mid \mathfrak{F}_{i-1})}{d^{(2)}(f, g)} \\ &= \frac{\beta_n^2(x)}{nd^{(2)}(f, g)} \sum_{i=1}^n \mathbb{E} \left[\left((\zeta(f, x) - \zeta(g, x)) \right. \right. \\ & \quad \left. \left. - \mathbb{E}[\zeta(f, x) - \zeta(g, x) \mid \mathcal{F}_{i-1}] \right)^2 \mid \mathfrak{F}_{i-1} \right]^2 \\ &\leq \frac{\beta_n^2(x)}{nd^{(2)}(f, g)} \sum_{i=1}^n 2\mathbb{E} \left[\left(\zeta(f, x) - \zeta(g, x) \right)^2 \mid \mathfrak{F}_{i-1} \right] \\ & \quad - 2\mathbb{E} \left\{ \left[\mathbb{E} \left[\left(\zeta(f, x) - \zeta(g, x) \right) \mid \mathfrak{F}_{i-1} \right] \right]^2 \right\} \\ &:= T_{1,n} + T_{2,n}. \end{aligned}$$

We first evaluate $T_{1,n}$. We have

$$\begin{aligned} T_{1,n} &\leq \frac{2\beta_n^2(x)}{nd^{(2)}(f, g)} \sum_{i=1}^n 2\mathbb{E} \left[\Delta_i^2(x) (f(Y_i) - g(Y_i))^2 \mid \mathfrak{F}_{i-1} \right] \\ & \quad + 2\mathbb{E} \left[\Delta_i^2(x) (\mathbb{G}(C_1, x) - \mathbb{G}(C_2, x))^2 \mid \mathfrak{F}_{i-1} \right] \\ &:= T_{1,n,1} + T_{1,n,2}. \end{aligned}$$

Using the fact that $\mathbb{E}(\Delta_1^2(x)) = \mathcal{O}(\phi(h))$ (in view of Lemma 1), the class of functions \mathcal{F} admits a constant envelope and $K(\cdot)$ is bounded and bounded

away from zero, one may get the following upper bound of the last equation, for some positive constant,

$$\begin{aligned}
 T_{1,n,1} &\leq \frac{C\sqrt{\phi(h)}}{d^{(2)}(f,g)} \mathbb{E} [\Delta_1(x) (f(Y_1) - g(Y_1))] \\
 &\leq \frac{C\sqrt{\phi(h)}}{d^{(2)}(f,g)} \mathbb{E} [\Delta_1(x)^2]^{1/2} \mathbb{E} [(f(Y_1) - g(Y_1))^2]^{1/2} \\
 &= \frac{C\sqrt{\phi(h)}}{G^2(\zeta)} \mathbb{E} [\Delta_1(x)^2]^{1/2} \\
 &= \mathcal{O}(\phi(h)) = o(1).
 \end{aligned}$$

Making use of similar arguments, we infer that

$$T_{1,n,2} = \frac{C\phi(h)^{3/2}}{d^{(2)}(f,g)} (\mathbb{E} [(f(Y) - g(Y))|X = x])^2 = \mathcal{O}(\phi(h)^{3/2}) = o(1).$$

We readily obtain that, $T_{1,n} = o(1)$. We have, by similar arguments to those used in the proof of the preceding statement, $T_{2,n} = o(1)$. Making use of Lindeberg conditions of the preceding proof and (32) combined with Theorem 1 of [4], we obtain, for given $\varepsilon > 0$ and $\gamma > 0$, there exists $\eta > 0$, such that

$$\limsup_{n \rightarrow \infty} \mathbb{P}^* \left\{ \sup_{d(C_1, C_2) \leq \eta} |\tilde{\nu}_n(C_1, x) - \tilde{\nu}_n(C_2, x)| \geq 5\gamma \right\} \leq 3\varepsilon. \tag{33}$$

Now the proof theorem is completed by combining this last equation with Theorem 3. □

7. Examples

Example 2 ([40]). On the Hilbert space \mathcal{E} equipped with the norm $\| \cdot \|$ associated to the inner product $\langle \cdot, \cdot \rangle$, consider the Hilbert autoregressive model of order one defined, for $n \geq 1$, by

$$X_n = \rho(X_{n-1}) + \epsilon_n,$$

where $(\epsilon_n)_{n \geq 1}$ is an i.i.d. sequence of Hilbert random variables such that ϵ_n is independent of X_{n-1} and $\mathbb{E}(\|\epsilon_n\|^2) < \infty$ and ρ is a functional operator on

\mathcal{E} . For $k \in \mathbb{N}^*$, consider the semi-metric d_k defined for any $(x, y) \in \mathcal{E}^2$, by

$$d_k(x, y) = \left(\sum_{j=1}^k \langle x - y, e_j \rangle^2 \right)^{\frac{1}{2}}. \tag{34}$$

Taking the semi-metric defined in the statement (34), observe that

$$F_x^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(d_k(x, X_i) \leq u \mid \mathcal{F}_{i-1}) = \mathbb{P}(d_k(x, \rho(X_i) + \epsilon_i) \leq u \mid \mathcal{F}_{i-1}).$$

Since we can write $\epsilon_i = \sum_{j=1}^{\infty} \epsilon_i^j e_j$ and for any $s \in \mathcal{E}$ $\rho(s) = \sum_{j=1}^{\infty} (\rho(s))_j e_j$, it follows that

$$\begin{aligned} F_{X_i \mid X_{i-1}=s}(u) &= \mathbb{P}(d_k(x, \rho(X_i) + \epsilon_i) \leq u \mid X_{i-1} = s) \\ &= \mathbb{P}\left(\sum_{j=1}^k \langle x_j - (\rho(s))_j + \epsilon_i^j, e_j \rangle^2 \leq u^2 \right) \\ &= \mathbb{P}(\|\bar{\epsilon}_i - (\bar{\rho}(s)) - x\|_{E_{cld}} \leq u) = \mathbb{P}(\bar{\epsilon}_i \in B_k((\bar{\rho}(s)) - x, u)), \end{aligned}$$

where $\bar{\epsilon}_i = (\epsilon_i^1, \dots, \epsilon_i^k)$, $\bar{\rho}(s) = ((\rho(s))_1, \dots, (\rho(s))_k)$ and $B_k(\bar{\rho}(s) - x, u)$ is the ball in \mathbb{R}^k of center $\bar{\rho}(s) - x$ and radius u . Denote by g the density function of $\bar{\epsilon}_i$. Clearly, we have

$$\begin{aligned} &F_{X_i \mid X_{i-1}=s}(u) \\ &= \int \dots \int_{B_k(\bar{\rho}(s)-x, u)} g(t_1, \dots, t_k) dt_1 \dots dt_k \\ &= \int \dots \int_{B_k(\bar{\rho}(s)-x, u)} |g(t_1, \dots, t_k) - g(\bar{\rho}(s) - x)| dt_1 \dots dt_k + C u^k g(\bar{\rho}(s) - x). \end{aligned}$$

When g is assumed to be a Lipschitz function of order 1 with a constant $C > 0$, we obtain

$$F_{X_i \mid X_{i-1}=s}(u) = C u^k g(\bar{\rho}(s) - x) + o(u^k).$$

Therefore,

$$F_x^{\mathcal{F}_{i-1}}(u) = F_{X_i \mid X_{i-1}=s}(u) = C u^k g(\bar{\rho}(X_{i-1}) - x) + o(u^k).$$

Example 3 ([40]). Let \mathcal{C} be a separate abstract space equipped with a semi-

distance. Consider the autoregressive model of order one defined, for any $i \geq 1$, by $X_i = \rho(X_{i-1}) + \epsilon_i$ where $\epsilon_i = \eta_i h$ with a real random variable η_i independent of X_{i-1} and $h \in \mathcal{C}$ and ρ is a functional operator on \mathcal{C} . For $(x, y) \in \mathcal{C}$ consider the semi-distance between x and y given by

$$d(x, y) = \left| \int (x(t) - y(t))dt \right|.$$

Observe, for any $u > 0$, that we have

$$F_x^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(d_k(x, X_i) \leq u \mid \mathcal{F}_{i-1}) = \mathbb{P}(d(x, X_i) \leq u \mid X_{i-1}).$$

Consequently, whenever $0 \neq \int h(t)dt < \infty$, we have

$$\begin{aligned} &F_{X_i|X_{i-1}=s}(u) \\ &= \mathbb{P}(d(x, X_i) \leq u \mid X_{i-1} = s) \\ &= \mathbb{P}\left(\left|\int x(t) - X_i(t)dt\right| \leq u \mid X_{i-1} = s\right) \\ &= \mathbb{P}\left(\left|\int x(t) - \rho(X_{i-1})(t) - \eta_i h(t)dt\right| \leq u \mid X_{i-1} = s\right) \\ &= \mathbb{P}\left(\left|\int x(t) - \rho(s)(t) - \eta_i h(t)dt\right| \leq u\right) \\ &= \mathbb{P}\left(\frac{-u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt} \leq \eta_i \leq \frac{u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right) \\ &= \Phi\left(\frac{u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right) - \Phi\left(\frac{-u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right), \end{aligned}$$

where Φ is the cumulative distribution function of η_i . Assuming now that $0 < \int h(t)dt < \infty$, $|\int x(t)dt| < \infty$ and $|\int \rho(s)(t)dt| < \infty$ for any $s \in \mathcal{C}$ and tacking Φ as the $\mathcal{N}(0, 1)$ cumulative distribution function, we obtain

$$F_{X_i|X_{i-1}=s}(u) = \frac{u}{\int h(t)dt} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2} \left(\frac{\int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right)^2\right) (1 + o(1)).$$

Thus

$$F_x^{\mathcal{F}_{i-1}}(u) = \frac{u}{\int h(t)dt} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2} \left(\frac{\int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right)^2\right) (1 + o(1)),$$

and the condition **(H1)**(ii) is satisfied with

$$\phi(u) = \frac{u}{\int h(t)dt} \sqrt{\frac{2}{\pi}}.$$

References

1. D. W. K. Andrews and D. Pollard, An introduction to functional central limit theorems for dependent stochastic processes *International Statistical Review / Revue Internationale de Statistique*, **62**(1994), No. 1, 119-132.
2. G. Aneiros, E. G. Bongiorno, R. Cao and P. Vieu, editors (2017). *Functional statistics and related fields. Selected papers based on the presentations at the 4th international workshop on functional and operational statistics*, IWFOS, Corunna, Spain, June 15-17, 2017. Cham: Springer.
3. M. A. Arcones, Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.*, **22**(1994), No. 4, 2242-2274.
4. J. Bae, D. Jun and S. Levental, The uniform CLT for martingale difference arrays under the uniformly integrable entropy, *Bull. Korean Math. Soc.*, **47**(2010), No. 1, 39-51.
5. R. F. Bass and R. Pyke, A strong law of large numbers for partial-sum processes indexed by sets, *Ann. Probab.*, **12**(1984), No. 1, 268-271.
6. E. G. Bongiorno, A. Goia, E. Salinelli and P. Vieu, . An overview of IWFOS'2014, In *Contributions in infinite-dimensional statistics and related topics*, pages 1-5(2014). Esculapio, Bologna.
7. D. Bosq, Linear processes in function spaces, volume 149 of *Lecture Notes in Statistics*. Springer-Verlag, New York. Theory and applications 2000.
8. D. Bosq and D. Blanke, *Inference and Prediction in Large Dimensions*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester; Dunod, Paris 2007.
9. S. Bouzebda and M. Chaouch, Uniform limit theorems for a class of conditional z -estimators when covariates are functions, *J. Multivariate Anal.*, **187**(2022), Paper accepted, 24.
10. S. Bouzebda and S. Didi, Additive regression model for stationary and ergodic continuous time processes, *Comm. Statist. Theory Methods*, **46**(2017a), No. 5, 2454-2493.
11. 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**(2017b), No. 3, 1367-1406.
12. 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** (2021), No. 3, 811-852.

13. S. Bouzebda and T. El-hadjali, Uniform convergence rate of the kernel regression estimator adaptive to intrinsic dimension in presence of censored data, *J. Nonparametr. Stat.*, **32**(2020), No. 4, 864-914.
14. 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** (2020), No. 2, 452-509.
15. S. Bouzebda, and B. Nemouchi, Weak-convergence of empirical conditional processes and conditional U -processes involving functional mixing data, (2021). Submitted.
16. 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, (2021) Submitted.
17. 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** (2015), No. 3, 163-199.
18. S. Bouzebda, M. Chaouch and N. Laïb , . Limiting law results for a class of conditional mode estimates for functional stationary ergodic data, *Math. Methods Statist.*, **25**(2016), No. 3, 168-195.
19. S. Bouzebda, F. Madani and Y. Souddi, Some asymptotic properties of the conditional set-indexed empirical process based on dependent functional data, *Int. J. Math. Stat.*, Submitted (2021), Accepted.
20. R. C. Bradley, *Introduction to strong mixing conditions*, Vol. 3. Kendrick Press, Heber City, UT 2007.
21. Y. S. Chow and H. Teicher, *Probability theory*, Springer Texts in Statistics. Springer-Verlag, New York, third edition 1997. Independence, interchangeability, martingales.
22. J. Dony and D. M. Mason, Uniform in bandwidth consistency of conditional U -statistics, *Bernoulli*, **14**(2008), No. 4, 1108-1133.
23. P. Doukhan, P. Massart and E. Rio, Invariance principles for absolutely regular empirical processes, *Ann. Inst. H. Poincaré Probab. Statist.*, **31** (1995), No. 2, 393-427.
24. R. M. Dudley, Central limit theorems for empirical measures. *Ann. Probab.*, **6** (1978), No. 6, 899-929.
25. R. M. Dudley, . A course on empirical processes, In *École d'été de probabilités de Saint-Flour, XII-1982*, volume 1097 of *Lecture Notes in Math.*, pages 1-142. Springer, Berlin 1984.
26. R. M. Dudley, *Uniform central limit theorems*, volume 63 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge 1999.
27. E. Eberlein, Weak convergence of partial sums of absolutely regular sequences, *Statist. Probab. Lett.*, **2** (1984), No. 5, 291-293.
28. F. Ferraty and P. Vieu, *Nonparametric functional data analysis*, Springer Series in Statistics. Springer, New York. Theory and practice 2006.

29. F. Ferraty, A. Mas and P. Vieu, . Nonparametric regression on functional data: inference and practical aspects, *Aust. N. Z. J. Stat.*, **49** (2007), No. 3, 267-286.
30. E. Giné and J. Zinn, Some limit theorems for empirical processes, *Ann. Probab.*, **12** (1984), No. 4, 929-998. With discussion.
31. L. Györfi, G. Morvai and S. J. Yakowitz, Limits to consistent on-line forecasting for ergodic time series, *IEEE Trans. Inform. Theory*, **44** (1998), No. 2, 886-892.
32. P. Hall, Asymptotic properties of integrated square error and cross-validation for kernel estimation of a regression function, *Z. Wahrsch. Verw. Gebiete*, **67** (1984), No. 2, 175-196.
33. P. Hall and C. C. Heyde, *Martingale limit theory and its application*, Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980.
34. W. Härdle and J. S. Marron, Optimal bandwidth selection in nonparametric regression function estimation, *Ann. Statist.*, **13** (1985), No. 4, 1465-1481.
35. J. Hoffmann-Jørgensen, *Stochastic processes on Polish spaces*, volume 39 of *Various Publications Series (Aarhus)*. Aarhus Universitet, Matematisk Institut, Aarhus, 1991.
36. L. Horváth and P. Kokoszka, *Inference for Functional Data with Applications*, Springer Series in Statistics. Springer, New York, 2012.
37. L. Horváth and B. S. Yandell, Asymptotics of conditional empirical processes, *J. Multivariate Anal.*, **26** (1988), No. 2, 184-206.
38. T. Hsing and R. Eubank, *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*, Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester 2015.
39. M. R. Kosorok, *Introduction to Empirical Processes and Semiparametric Inference*, Springer Series in Statistics. Springer, New York, 2008.
40. N. Laib and D. Louani, Nonparametric kernel regression estimation for functional stationary ergodic data: asymptotic properties, *J. Multivariate Anal.*, **101** (2010), No. 10, 2266-2281.
41. L. Le Cam, A remark on empirical measures, In *A Festschrift for Erich L. Lehmann*, Wadsworth Statist./Probab. Ser., pages 305-327(1983). Wadsworth, Belmont, CA.
42. E. A. Nadaraja, On a regression estimate. *Teor. Verogatnost. i Primenen.*, **9**(1964), 157-159.
43. A. Nobel and A. Dembo, A note on uniform laws of averages for dependent processes, *Statist. Probab. Lett.*, **17** (1993), No. 3, 169-172.
44. D. Pollard, A central limit theorem for empirical processes, *J. Austral. Math. Soc. Ser. A*, **33**(1982), No. 2 , 235-248.
45. W. Polonik and Q. Yao, Set-indexed conditional empirical and quantile processes based on dependent data, *J. Multivariate Anal.*, **80** (2002), No. 2, 234-255.

46. D. V. Poryvaĭ, An invariance principle for conditional empirical processes formed by dependent random variables, *Izv. Ross. Akad. Nauk Ser. Mat.*, **69** (2005), No. 4, 129-148.
47. M. Rachdi and P. Vieu, . Nonparametric regression for functional data: automatic smoothing parameter selection, *J. Statist. Plann. Inference*, **137** (2007), No. 9, 2784-2801.
48. J. O. Ramsay and B. W. Silverman, *Functional data analysis*, Springer Series in Statistics. Springer, New York, second edition 2005a.
49. J. O. Ramsay and B. W. Silverman, *Functional data analysis*, Springer Series in Statistics. Springer, New York, second edition 2005b.
50. M. Rosenblatt, Uniform ergodicity and strong mixing, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **24**(1972), 79-84.
51. H. L. Shang, Bayesian bandwidth estimation for a functional nonparametric regression model with mixed types of regressors and unknown error density, *J. Nonparametr. Stat.*, **26** (2014), No. 3, 599-615.
52. J. Q. Shi and T. Choi, *Gaussian Process Regression Analysis for Functional Data*, CRC Press, Boca Raton, FL2011.
53. W. Stute, Conditional empirical processes, *Ann. Statist.*, **14** (1986a), No. 2, 638-647.
54. W. Stute, On almost sure convergence of conditional empirical distribution functions, *Ann. Probab.*, **14** (1986b), No. 3, 891-901.
55. A. W. van der Vaart and J. A. Wellner, *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York. With applications to statistics 1996.
56. V. N. Vapnik and A. J. Červonenkis, The uniform convergence of frequencies of the appearance of events to their probabilities, *Teor. Verojatnost. i Primenen.*, **16**(1971), 264-279.
57. G. S. Watson, Smooth regression analysis, *Sankhyā Ser. A*, **26**(1964), 359-372.
58. .-i. Yoshihara, Conditional empirical processes defined by ϕ -mixing sequences. *Comput. Math. Appl.*, **19** (1990), No. 1, 149-158.
59. B. Yu, Rates of convergence for empirical processes of stationary mixing sequences, *Ann. Probab.*, **22** (1994), No. 1, 94-116.
60. J.-T. Zhang, Analysis of variance for functional data, volume 127 of *Monographs on Statistics and Applied Probability*, CRC Press, Boca Raton, FL 2014.