Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 9(2): 4427–4550.
DOI: 10.3934/math.2024218
Received: 27 November 2023
Revised: 27 December 2023
Accepted: 05 January 2024
Published: 18 January 2024

Research article

Uniform in number of neighbors consistency and weak convergence of kNN empirical conditional processes and kNN conditional U-processes involving functional mixing data*

Salim Bouzebda* **and Amel Nezzal**

LMAC (Laboratory of Applied Mathematics of Compiègne), Université de technologie de Compiègne, France

* **Correspondence:** Email: salim.bouzebda@utc.fr.

Abstract: U-statistics represent a fundamental class of statistics arising from modeling quantities of interest defined by multi-subject responses. U-statistics generalize the empirical mean of a random variable X to sums over every m-tuple of distinct observations of X. Stute [182] introduced a class of so-called conditional U-statistics, which may be viewed as a generalization of the Nadaraya-Watson estimates of a regression function. Stute proved their strong pointwise consistency to: $r^{(m)}(\varphi, \mathbf{t}) := \mathbb{E}[\varphi(Y_1, \ldots, Y_m)|(X_1, \ldots, X_m) = \mathbf{t}]$, for $\mathbf{t} \in \mathcal{X}^m$. In this paper, we are mainly interested in the study of the kNN conditional U-processes in a functional mixing data framework. More precisely, we investigate the weak convergence of the conditional empirical process indexed by a suitable class of functions and of the kNN conditional U-processes when the explicative variable is functional. We treat the uniform central limit theorem in both cases when the class of functions is bounded or unbounded satisfying some moment conditions. The second main contribution of this study is the establishment of a sharp almost complete Uniform consistency in the Number of Neighbors of the constructed estimator. Such a result allows the number of neighbors to vary within a complete range for which the estimator is consistent. Consequently, it represents an interesting guideline in practice to select the optimal bandwidth in nonparametric functional data analysis. These results are proved under some standard structural conditions on the Vapnik-Chervonenkis classes of functions and some mild conditions on the model. The theoretical results established in this paper are (or will be) key tools for further functional data analysis developments. Potential applications include the set indexed conditional U-statistics, Kendall rank correlation coefficient, the discrimination problems and the time series prediction from a continuous set of past values.

Keywords: conditional U-statistic; functional data analysis; functional regression; Kolmogorov's entropy; small ball probability; kNN Kernel-type estimators; empirical processes; U-processes; weak convergence; VC-classes

*Both authors contributed equally to this work.

Mathematics Subject Classification: 60F05, 60F15, 62E20, 62G05, 62G07, 62G08, 62G20, 62G35

1. Introduction and motivations

U-statistics were first introduced by [115] in connection with unbiased estimators, following initial work by [106]. In brief, U-statistics of order m and kernel h based on a sequence $\{X_i\}_{i=1}^{\infty}$ of random variables with values in a measurable space $(S, \mathfrak{S})$ and a measurable function $f : S^m \to \mathbb{R}$ are given by

$$U_n(h) = \frac{(n-m)!}{n!} \sum_{(i_1,\ldots,i_m)\in I_n^m} h\left(X_{i_1}, \ldots, X_{i_m}\right), \quad n \geq m,$$

where

$$I_n^m = \left\{(i_1, \ldots, i_m) : i_j \in \mathbb{N}, 1 \leq i_j \leq n, i_j \neq i_k \text{ if } j \neq k\right\}.$$

$U_n(h)$ is the nonparametric uniformly minimum variance estimator of $\theta = \mathbb{E}\left(h\left(X_1, \ldots, X_m\right)\right)$. It is the minimizer with respect to α of

$$\sum_{1\leqslant i_1<\cdots<i_m\leqslant n} \left(h\left(X_{i_1}, \ldots, X_{i_m}\right) - \alpha\right)^2.$$

Empirical variance, Gini's mean difference, and Kendall's rank correlation coefficient are common examples of estimators based on U-statistics. The Wilcoxon signed rank test for the hypothesis of the location at zero is a classical test based on U-statistics, as discussed by [190] in Example 12.4. Asymptotic results for the case of independent and identically distributed underlying random variables were first provided by [115], who also referred to related work by [57, 83, 172, 176]. Similar results were obtained for V-statistics by [94, 198]. Extensive literature on the theory of U-statistics has been developed, as reviewed by [10, 138, 176], and others. A detailed review and major historical developments in this field can be found in the book by [20]. U-processes are sets of U-statistics indexed by a family of kernels, which can be viewed as infinite-dimensional variants of U-statistics with a single kernel or as stochastic processes that are nonlinear extensions of empirical processes. U-processes have been applied to solve complex statistical problems such as density estimation, nonparametric regression tests, and goodness-of-fit tests. Considering a large group of statistics instead of a single statistic is more statistically interesting, and ideas from the theory of empirical processes can be used to construct limit or approximation theorems for U-processes. However, obtaining results for U-processes is not easy and requires significant effort and distinct methodologies. Generalizing from empirical processes to U-processes is particularly difficult, especially in the stationary setting. U-processes appear in statistics in many instances, such as the components of higher-order terms in von Mises expansions, and play a role in analyzing estimators (including function estimators) with varying degrees of smoothness. For instance, the product limit estimator for truncated data is analyzed in [183] using a.s. uniform bounds for $\mathbb{P}$-canonical U-processes. In addition, [11] introduce two new tests for normality based on U-processes, while [173] use weighted L_1-distances between the standard normal density and local U-statistics based on standardized observations to propose new tests for normality, utilizing the results of [101]. Moreover, in [122], the median-of-means approach, which is based on

U-statistics, is introduced to estimate the mean of multivariate functions in case of possibly heavy-tailed distributions. U-processes play a significant role in various statistical applications, including testing for qualitative features of functions in nonparametric statistics [1, 100], cross-validation for density estimation [157], and establishing limiting distributions of M-estimators [10, 68, 177]. [10] provide necessary and sufficient conditions for the law of large numbers and sufficient conditions for the central limit theorem for U-processes. For further references on U-statistics and U-processes, interested readers may refer to [28, 40, 46, 47, 48, 54, 138, 179], while a comprehensive insight into the U-processes theory is provided by [68]. U-statistics are also naturally found in other contexts, such as the theory of random graphs, where they count occurrences of specific subgraphs like triangles, as presented in [120]. In machine learning, U-statistics arise naturally in various problems such as clustering, image recognition, ranking, and learning on graphs, where natural risk estimates take the form of U-statistics, as discussed in [63]. For instance, the empirical ranking error of any given prediction rule is a U-statistic of order 2, as stated in [62]. For U-statistics with random kernels of diverging orders, readers may refer to [97, 112, 178, 180]. Infinite-order U-statistics are also useful for constructing simultaneous prediction intervals that quantify the uncertainty of ensemble methods like subbagging and random forests, as presented in [161]. The MeanNN approach estimation for differential entropy, introduced by [90], is a particular application of the U-statistic. Additionally, [143] proposed a new test statistic for goodness-of-fit tests using U-statistics. Moreover, [65] have explored a model-free approach for clustering and classifying genetic data based on U-statistics, leading to alternative ways of addressing genetic problems. Their motivation was based on the versatility and adaptability of U-statistics to various genetic problems and different data types. [140] proposed using the U-statistics, in a natural way, for analyzing random compressed sensing matrices in the non-asymptotic regime. Extending the above exploration to conditional U-processes is practically useful and technically more challenging.

We first introduce Stute's estimators. Let us consider regular sequence of random elements $\{(\mathbf{X}_i, Y_i), i \in \mathbb{N}^*\}$ with $\mathbf{X}_i \in \mathbb{R}^d$ and $Y_i \in \mathcal{Y}$ some polish space and $\mathbb{N}^* = \mathbb{N}\backslash\{0\}$. Let $\varphi : \mathcal{Y}^m \to \mathbb{R}$ be a measurable function. In this paper, we are primarily concerned with the estimation of the conditional expectation, or regression function, for $\mathbf{t} \in \mathbb{R}^{dm}$,

$$r^{(m)}(\varphi, \mathbf{t}) \quad = \quad \mathbb{E}\left(\varphi(Y_1, \ldots, Y_m) \mid (\mathbf{X}_1, \ldots, \mathbf{X}_m) = \mathbf{t}\right), \tag{1.1}$$

whenever it exists, i.e.,

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

We now introduce a kernel function $K : \mathbb{R}^d \to \mathbb{R}$ with support contained in $[-B, B]^d$, $B > 0$, satisfying:

$$\sup_{\mathbf{x}\in\mathbb{R}^d} |K(\mathbf{x})| =: \kappa < \infty \quad \text{and} \quad \int K(\mathbf{x})d\mathbf{x} = 1. \tag{1.2}$$

[182] introduced a class of estimators for $r^{(m)}(\varphi, \mathbf{t})$, called conditional U-statistics, which is defined for each $\mathbf{t} \in \mathbb{R}^{dm}$ to be :

$$\widehat{r}_n^{(m)}(\varphi, \mathbf{t}; h_K) = \frac{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} \varphi(Y_{i_1}, \ldots, Y_{i_m}) K\left(\frac{\mathbf{t}_1 - \mathbf{X}_{i_1}}{h_K}\right) \cdots K\left(\frac{\mathbf{t}_m - \mathbf{X}_{i_m}}{h_K}\right)}{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{\mathbf{t}_1 - \mathbf{X}_{i_1}}{h_K}\right) \cdots K\left(\frac{\mathbf{t}_m - \mathbf{X}_{i_m}}{h_K}\right)}, \tag{1.3}$$

where

$$I(m,n) = \left\{ \mathbf{i} = (i_1, \ldots, i_m) : 1 \leq i_j \leq n \quad \text{and} \quad i_j \neq i_r \quad \text{if} \quad j \neq r \right\},$$

is the set of all m-tuples of different integers between 1 and n and $\{h_K := h_n\}_{n\geq 1}$ is a sequence of positive constants converging to zero at the rate $nh_K^m \to \infty$. In the particular case $m = 1$, the $r^{(m)}(\varphi, t)$ is reduced to

$$r^{(1)}(\varphi, t) = \mathbb{E}(\varphi(\mathbf{Y})|\mathbf{X} = t),$$

and Stute's estimator becomes the Nadaraya-Watson [153, 200] estimator of $r^{(1)}(\varphi, t)$, refer to [50] for details. The work of [175] was devoted to estimating the rate of the uniform convergence in $\mathbf{t}$ of $\widehat{r}_n^{(m)}(\varphi, \mathbf{t}; h_K)$ to $r^{(m)}(\varphi, \mathbf{t})$. In [165], the limit distributions of $\widehat{r}_n^{(m)}(\varphi, \mathbf{t}; h_K)$ are analyzed and compared to those derived by Stute. Under proper mixing settings, [111] extended the results of [182] to weakly dependent data and applied their findings to validate the Bayes risk consistency of the corresponding discrimination rules. As alternatives to the conventional kernel-type estimators, [186] offered symmetrized closest neighbor conditional U-statistics. [98] evaluated the functional conditional U-statistic and determined its finite-dimensional asymptotic normality. Nonparametric estimate of the conditional U-statistics in a functional data context has gotten comparatively little attention despite the subject's significance. Recent developments are described in [42, 44, 45, 52, 56], in which the authors examine challenges associated with the uniform in bandwidth consistency in general settings. [119] evaluated the test of independence in the functional framework using the Kendall statistics, which may be viewed as special cases of the U-statistics. [14] introduced a comprehensive framework for clustering within multiple groups, employing a U-statistics-based approach specifically designed for high-dimensional datasets. This method classifies data into three groups while evaluating the significance of these partitions. In a related context, [128] focused on dimension-agnostic inference, devising methods whose validity remains independent of assumptions regarding dimension versus sample size. Their approach utilized variational representations of existing test statistics, incorporating sample splitting and self-normalization to yield a refined test statistic with a Gaussian limiting distribution. This modification of degenerate U-statistics involved dropping diagonal blocks and retaining off-diagonal blocks. Exploring further, [59] delved into U-statistics-based empirical risk minimization, while [121] examined asymmetric U-statistics based on a stationary sequence of m-dependent variables, with applications motivated by pattern matching in random strings and permutations. Additionally, [188] developed innovative U-statistics considering left truncation and right censoring. As an application, they proposed a straightforward non-parametric test for assessing the independence between time to failure and the cause of failure in competing risks, particularly when observations are subject to left truncation and right censoring. In a different context, [136] investigated the quadruplet U-statistic, offering applications in statistical inference for network analysis. It will be interesting to find connection of the U-statistics with the problems investigated in [187, 201, 204, 205]. The extension of the preceding investigation to conditional empirical U-processes is both practically beneficial and technically difficult.

Recently, there has been a growing interest in regression models in which the response variable is real-valued and the explanatory variable is represented by smooth functions that vary arbitrarily between repeated observations or measurements. This form of data, known as functional data, appears in numerous disciplines, such as climatology (hourly concentration of pollutants), medicine (the knee angle of children as functions of time), economics, linguistics, etc. Functional time series are commonly encountered in practice, for example, when a long continuous time process is divided into smaller natural units, such as days. In this instance, every intraday curve is a functional random variable. This paper focuses mostly on the instance of functional data and the theory of the U-processes. We give an excerpt from [6]: Functional data analysis (FDA) is a branch of statistics that focuses on the study of variables having an unlimited number of dimensions, such as curves, sets, and images. It has had spectacular growth over the past two decades, fueled partly by technological advances that have led to the "Big Data" revolution. FDA is today one of the most active and significant disciplines of research in data science, despite its reputation at the turn of the century as a fairly obscure area of study. The reader is recommended to the works of [167, 168], [91] for an introduction to this field. These sources offer a variety of case studies in numerous fields, including criminology, economics, archaeology, and neurophysiology, as well as basic analysis techniques. It should be noted that the extension of probability theory to random variables with values in normed vector spaces (such as Banach and Hilbert spaces), as well as extensions of certain classical asymptotic limit theorems, predates the recent literature on functional data; the reader is referred to [7]. [99] examined density and mode estimates for data occupying a normed vector space. This study examines the topic of the curse of dimensionality for functional data and proposes solutions to the problem. In the context of regression estimation, [91] examined nonparametric models. We may also refer to [21, 116]. [130] provided a nice mix of foundational material, accessible theory, and practical examples. Recently, the contemporary theory was used in the analysis of functional data. [93], who has provided the consistency rates of several functionals of the conditional distribution, such as the regression function, the conditional cumulative distribution, the conditional density, and others, uniformly over a subset of the explanatory variable. [125] established the consistency rates for some functionals nonparametric models, such as the regression function, the conditional distribution, the conditional density, and the conditional hazard function, uniformly in bandwidth (UIB consistency). [35] extended these results to the ergodic setting. [12] examined the issue of local linear estimation of the regression function when the regressor is functional and demonstrated strong convergence (with rates) uniformly in bandwidth parameters. [141] examined the k-nearest neighbours (kNN) estimate of the nonparametric regression model for strong mixing functional time series data and demonstrated the uniform nearly complete convergence rate of the kNN estimator under several moderate conditions. [30] provided several limiting law results for the conditional mode in the functional setting for ergodic data; for more recent references, see [3, 4, 5, 53, 55, 151].

The primary objective of this study is to examine a generic framework and characterize the weak convergence and the uniform convergence of kNN conditional U-processes based on a regular sequence of random functions. This is motivated by the fact that the k-NN method is a fundamental statistical method presenting several advantages. Recall that the k-NN method considers the k neighbors of X_i nearest to x with respect to some distance $d(\cdot,\cdot)$. Although the local bandwidth of the k-NN is random and depends on the data X_i respecting the local structure of the data, which is essential in the infinite dimension. Historically, the k-NN was first introduced by [95]-see also [96]-in the context of nonparametric discrimination, and further investigated by [144], for more details, we refer to [17]. It is commonly used in practice (see [91]) and is simple to handle because the user has only one parameter to control the number k of nearest neighbors, valued in a finite set. In addition, it allows us to build a neighbor adapted to the data at any point. The k-NN method is widely studied if the explanatory variable is an element of a finite-dimensional space, for instance, see [16, 64, 77, 104, 145]. In an infinite dimensional space, i.e., a functional framework, there are three different approaches for the

k-NN regression estimation. The first, published by [135], examines a k-NN kernel estimate when the functional variable is an element of a separable Hilbert space H. In this approach, [135] established a weak consistency result. The strategy of [135] is to reduce the infinite dimension of H by using a projection on a finite dimension subspace by considering only the first m coefficients of an expansion of X in an orthonormal system of H and then applying the multivariate techniques on the projected data to perform the k-NN regression. The second approach is based on the k-NN procedure and the functional local linear estimation, the consistency with the convergence rate is obtained in [60] and [142].

More precisely, in this paper, we are interested in establishing the a.co uniform consistency and the a.co uniform in the number of neighbors (UINN) consistency of the nonparametric functional regression estimator and also the functional conditional U-processes (statistics). [157] were the first to introduce the notion of uniform in bandwidth consistency for kernel density estimators and they applied empirical process methods in their study. This is motivated by a series of papers, among many others, [15, 25, 33, 36, 37, 38, 42, 44, 51, 52, 54, 71, 78, 87, 88] the authors established uniform in bandwidth (UIB) consistency results for such estimators in the i.i.d. finite-dimensional setting, where h_n varies within suitably chosen intervals indexed by n. In the FDA, several authors have been interested in studying non-parametric functional estimators. For example, [93] provided the consistency rates of some functionals of the conditional distribution, including the regression function, the conditional cumulative distribution, the conditional density, and some others, uniformly over a certain subset of the explicative variable. [125] established the uniform consistency rate for some conditional models, including the regression function, the conditional distribution, the conditional density, and the conditional hazard function. The last mentioned paper is extended by [42]. [124] and [3] established the almost complete convergence of the k-nearest neighbors (k-NN) estimators, which are uniform in the number of neighbors, under some classical assumptions on the kernel and on the small ball probabilities of the functional variable in connection with the entropy condition controlling the space complexity. [12] considered the problem of local linear estimation of the regression function when the covariate is functional and proved the strong uniform-in-bandwidth (UIB) convergence. [141] investigated the k-NN estimation of the nonparametric regression model for strong mixing functional time series data and established the uniform a.co convergence rate of the k-NN estimator under some mild conditions. [158] stated some new uniform asymptotic results for kernel estimates in the functional single-index model. Most of this literature focuses on UIB or UINN consistency or on uniform consistency on some functional subset but never the both together, which was investigated in [44] in the independent framework. We aim to fill this gap in the literature by combining results from the FDA and the empirical processes theory in the dependent setting. The second problem for the weak convergence that we investigate is not simple, and the main merits of our contribution are the control of the asymptotic equi-continuity under minimal conditions in this general setting, which constitutes a fundamentally unresolved open problem in the literature. We intend to fill this gap in the literature by integrating the results of [8] and [27] with the strategies described in [147] and [43] for handling functional data. But, as will be demonstrated in the following section, the challenge requires much more than "just" merging concepts from the current outcomes. In reality, intricate mathematical derivations will be necessary to deal with the typical functional data in our framework. This necessitates the effective application of large sample theory tools, which were established for dependent empirical processes and for which we have used results from the work of

[8, 27, 43]. Even with i.i.d. functional data, no weak convergence for the kNN conditional U-processes has been proven up to the present.

1.1. Paper contribution

The current paper looks into the challenges high-dimensional functional data presents and presents important findings that are applicable to high-dimensional data models. This research expands the classical kernel estimator beyond samples that are independent and identically distributed to include stationary random processes. In addition, the main focus is on the kNN kernel estimator. It specifically examines four fundamental aspects that pertain to the kNN kernel estimator for regression and its uniform consistency. The study investigates the UINN and UIB scenarios to guarantee consistent results within the domain of functional regression (the UIB in Theorem 3.3 and the UINN in Theorem 3.1). The outcomes include the prediction of relative error, which enhances the overall comprehension of the kNN kernel estimator's performance (the UIB in Corollary 3.6 and and the UINN in Corollary 3.5). The paper includes the kNN method's for functional conditional U-statistics with consistent results, thereby extending the application area of these estimators. The study additionally investigates the uniform consistency and UINN consistency of functional conditional U-statistics, offering a comprehensive assessment of their limiting behaviour (the UIB in Theorems 3.8–3.10 and Corollary 3.12 and the UINN in Theorems 3.14, 3.16 and Corollary 3.18). In addition, the study makes an advanced contribution to the field by establishing a uniform central limit theorem for classes of functions that, subject to specific moment conditions, are either bounded or unbounded (for the conditional process the Normality is provided in Theorem 4.1 and the equicontinuity in Theorem 4.2; for the conditional U-process the Normality is provided in Theorems 4.5, 4.6 and the equicontinuity in Theorem 4.7). The process of establishing our main results uses advanced methodologies, including the k-nearest neighbours (kNN) method, covering number, small-ball probability, the Hoeffding decomposition, the decoupling methods, and the modern theory of the empirical process indexed by functions. At this stage, we mention that the Hoeffding decomposition cannot be used directly in our setting and needs some intricate preparation. All these results are established under fairly general conditions on function classes and underlying distributions, the majority of which are derived from prior works, thus guaranteeing their feasibility. The obtained results have the potential to be utilised in numerous statistical domains, such as time series prediction, set-indexed conditional U-statistics, and the Kendall rank correlation coefficient. Key technical tools in the proofs are the maximal moment inequalities for U-processes, and [84]'s results on β-mixing.

1.2. Paper Organization

The layout of the present article is as follows. Section 2 is devoted to introducing the functional framework and the definitions that we need in our work, we give the assumptions used in our asymptotic analysis with a short discussion. Section 3 is devoted to the strong uniform convergence with rate. Section 4.1 provides the weak convergence of empirical processes in the functional framework. Section 4.2 gives the main results of the paper concerning the uniform TCL for the conditional U-processes. In Section 5, we collect some potential applications, including the set indexed conditional U-statistics in Section 5.1, Kendall rank correlation coefficient in Section 5.2, the discrimination problems in Section 5.3 and the time series prediction from a continuous set of past values in Section 5.4. We discuss a bandwidth choice for practical use in Section 6. Some concluding remarks and possible future developments are relegated to Section 7. To prevent interrupting the flow of the presentation, all proofs, based upon modern empirical process theory, are gathered in Section 8. Due to the lengthiness of the proofs, we limit ourselves to the most important arguments. A few relevant technical results are given in the Appendix.

2. The functional framework

2.1. Generality on the model

Let $\{(X_i, Y_i) : i \geq 1\}$ be a sequence of stationary[†] random copies of the random vector [rv] (X, Y), where X takes its values in some abstract space $\mathcal{X}$ and Y in the abstract space $\mathcal{Y}$. Suppose that $\mathcal{X}$ is endowed with a semi-metric $d(\cdot,\cdot)$[‡] defining a topology to measure the proximity between two elements of $\mathcal{X}$ and which is disconnected from the definition of X to avoid measurability problems. We are mainly interested in establishing the weak convergence of the conditional U-process based on the following U-statistic in the k-NN setting introduced in [44] by

$$\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t})) = \frac{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} \varphi(Y_{i_1},\ldots,Y_{i_m}) K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}, \tag{2.1}$$

as an estimator for the multivariate regression function

$$r^{(m)}(\varphi, \mathbf{t}) := \mathbb{E}\left(\varphi\left(Y_1, \ldots, Y_m\right) \mid X_1, \ldots, X_m = \mathbf{t}\right), \tag{2.2}$$

where

$$\widetilde{K}(\mathbf{v}) = \prod_{i=1}^{m} K(v_i), \;\; \mathbf{v} = (v_1, \ldots, v_m) \in \mathbb{R}_+^m, \tag{2.3}$$

and $\varphi : \mathcal{Y}^m = \mathcal{Y} \times \ldots \times \mathcal{Y} \longrightarrow \mathbb{R}$ is a symmetric measurable function belonging to some class of functions $\mathscr{F}_m$, and $\mathbf{h}_{n,k}(\mathbf{t}) = (H_{n,k}(t_1), \ldots, H_{n,k}(t_m))$ is a vector of positive random variables that depend on $(X_1, \ldots, X_n)$ such that, for all $\mathbf{x} = (x_1, \ldots, x_m) \in \mathcal{X}^m$ and $j = 1, \ldots, m$:

$$H_{n,k}(x_j) = \min\left\{h \in \mathbb{R}^+ : \sum_{i=1}^{n} \mathbb{1}_{B(x_j,h)}(X_i) = k\right\}, \tag{2.4}$$

where $B(t, r) = \{z \in \mathcal{X} : d(z, t) \leqslant r\}$ is a ball in $\mathcal{X}$ with the center $t \in \mathcal{X}$ and radius r, and $\mathbb{1}_A$ is the indicator function of the set A. In fact, this k-NN estimate can be considered as an extension to the

[†]In the case of the Hilbert space valued elements not necessarily strictly stationary is needed, a second order stationarity suffices. An Hilbert space valued sequence $\{X_t\}_{t\in\mathbb{Z}}$ is second-order (or weakly) stationary if $\mathbb{E}\|X_t\|^2 < \infty, \mathbb{E}X_t = \mu$, and

$$\mathbb{E}\left((X_s - \mu) \otimes (X_t - \mu)\right) = \mathbb{E}\left((X_{s-t} - \mu) \otimes (X_0 - \mu)\right),$$

for all $s, t \in \mathbb{Z}$. We say that $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationary if the joint distribution of $\left\{X_{t_1}, \ldots, X_{t_n}\right\}$ and the joint distribution of $\left\{X_{t_1+h}, \ldots, X_{t_n+h}\right\}$ coincide, for all $t_1, \ldots, t_n \in \mathbb{Z}, n \geq 1$, and $h \geq 1$.

[‡]A semi-metric (sometimes called pseudo-metric) $d(\cdot,\cdot)$ is a metric which allows $d(x_1, x_2) = 0$ for some $x_1 \neq x_2$.

random and locally adaptive neighbor of the functional conditional U-statistics estimate of $r^{(m)}(\varphi, \mathbf{t})$ defined for all $t \in \mathcal{X}^m$ as :

$$\widehat{r}_n^{(m)}(\varphi, \mathbf{t}, h_K(\mathbf{t})) = \frac{\sum_{(i_1,\ldots,i_m)\in I(m,n)} \varphi(Y_{i_1}, \ldots, Y_{i_m}) K\left(\frac{d(t_1, X_{i_1})}{h_K(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{h_K(t_m)}\right)}{\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{d(t_1, X_{i_1})}{h_K(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{h_K(t_m)}\right)}, \tag{2.5}$$

where $h_K(\mathbf{t}) = (h_{K,n}(t_1), \ldots, h_{K,n}(t_m)) =: (h_K(t_1), \ldots, h_K(t_m))$ are positive real numbers decreasing to zero as n goes to infinity. At this stage, we highlight that kernel estimation is popular since the classical Akaike-Parzen-Rosenblatt kernel density estimation [refer to [2, 160, 171]]. However, the first appearance of kernel estimators is likely to be [95]: as the original technical report is difficult to find, it has been re-published as [96]. [160] has shown, under some assumptions on $K(\cdot)$, that $f_n(\cdot)$ is an asymptotically unbiased and consistent estimator for $f(\cdot)$ whenever $h_n \rightarrow 0, nh_n^d \rightarrow \infty$ and $\tilde{\mathbf{x}}$ is a continuity point of $f(\cdot)$. Under some additional assumptions on $f(\cdot)$ and h_n, he obtained an asymptotic normality result, too. The kernel estimators have been extensively studied in the literature, see, e.g., [28, 31, 34, 49, 73, 75, 76, 85, 86, 108, 154, 174, 199] and the references therein. The k-NN method is a fundamental statistical tool with various advantages. Generally, the procedure is computationally efficient and requires minimal parameter adjustment. In addition, the k-NN approaches are nonparametric, which allows them to automatically adapt to any continuous underlying distributions without relying on any particular models. The k-NN techniques were proven consistent for various significant statistical problems, including density estimation, classification, and regression, provided that a suitable k is chosen. It should be noted that since our objective is to generalize the results obtained for the estimator defined in (2.5), and given the fact that one of the main differences is that the smoothing parameter, $\mathbf{h}_{n,k}(\mathbf{t})$ is a vector of random variables instead of a univariate parameter h_K our first course of action would be to extend the results of [42, 43] to the multivariate setting. First, we need to introduce some notation. Let $\mathscr{F}_m = \{\varphi : \mathcal{Y}^m \longrightarrow \mathbb{R}\}$ denote a pointwise measurable class of real-valued symmetric functions on $\mathcal{Y}^m$ with a measurable envelope function

$$F(\mathbf{y}) \geq \sup_{\varphi\in\mathscr{F}_m} |\varphi(\mathbf{y})|, \text{ for } \mathbf{y} \in \mathcal{Y}^m. \tag{2.6}$$

For a kernel function $K(\cdot)$ and a subset $S_{\mathcal{X}} \subset \mathcal{X}$, we define the pointwise measurable class of functions, for $1 \leq m \leq n$:

$$\mathscr{K}^m := \left\{(x_1, \ldots, x_m) \mapsto \prod_{i=1}^{m} K\left(\frac{d(x_i, t_i)}{h_i}\right), \ (h_1, \ldots, h_m) \in \mathbb{R}_+^m \setminus \{0\} \text{ and } (\mathbf{t}_1, \ldots, \mathbf{t}_m) \in S_{\mathcal{X}}^m\right\}.$$

Statistical observations are not always independent but are often close to being so. Dependence may lead to severe repercussions on statistical inference if it is not taken into consideration. The notion of mixing quantifies how close to independence a sequence of random variables is, allowing us to extend standard results for independent sequences to weakly dependent or mixing sequences. Let us specify the dependence that will be the focus of this study. Let $\mathbf{Z}_1, \mathbf{Z}_2, \ldots$ be a stationary sequence of random variables on some probability space $(\Omega, \mathcal{D}, \mathbb{P})$ and let σ_i^j be the σ-field generated by $\mathbf{Z}_i, \ldots, \mathbf{Z}_j$, for

$i, j \geqslant 1$. The sequence $\mathbf{Z}_1, \mathbf{Z}_2, \ldots$ is said β-mixing or absolute regular, refer to [170, 197], if :

$$\beta(s) := \mathbb{E} \sup_{l \geqslant 1} \left\{ \left| \mathbb{P}\left(A | \sigma_1^l\right) - \mathbb{P}(A) \right| : A \in \sigma_{l+s}^{\infty} \right\} \longrightarrow 0 \text{ as } s \to \infty.$$

It should be noted that [118] obtained a complete description of stationary Gaussian processes satisfying the last property. Throughout the sequel, we assume tacitly that the sequence of random elements $\{(X_i, Y_i), i \in \mathbb{N}^*\}$ is absolutely regular. The Markov chains, for instance, are β-mixing under the milder Harris recurrence condition if the underlying space is finite [19, 46, 66, 180]. We also need to introduce some concepts that are related to the topological structure of functional spaces. First, we define the small-ball probability for a fixed $t \in \mathcal{X}$ and for all $r > 0$ by

$$\mathbb{P}(X \in B(t, r)) =: \phi_t(r), \tag{2.7}$$

this notion is widely used in nonparametric functional data analysis to avoid introducing density assumptions on the functional variable X and address the issues associated with the infinite-dimensional nature of the functional spaces. At this point, we can refer to [91, 99, 147]. We also need to deal with the VC-subgraph classes ("VC" for Vapnik and Chervonenkis, for instance, see [132, 193, 194]).

Definition 2.1. *A class of subsets C on a set C is called a VC-class if there exists a polynomial $\mathfrak{P}(\cdot)$ such that, for every set of N points in C, the class C picks out at most $\mathfrak{P}(N)$ distinct subsets.*

Definition 2.2. *A class of functions $\mathscr{F}$ is called a VC-subgraph class if the graphs of the functions in $\mathscr{F}$ form a V-C class of sets, that is, if we define the subgraph of a real-valued function f on S as the following subset $\mathcal{G}_f$ on $S \times \mathbb{R}$*

$$\mathcal{G}_f = \{(s, t) : 0 \leq t \leq f(s) \quad or \quad f(s) \leq t \leq 0\},$$

the class $\{\mathcal{G}_f : f \in \mathscr{F}\}$ is a VC-class of sets on $S \times \mathbb{R}$. Informally, a VC-class of functions is characterized by having a polynomial covering number (the minimal number of required functions to make a covering on the entire class of functions).

Definition 2.3. *Let $\mathcal{S}_{\mathcal{E}}$ be a subset of a semi-metric space $\mathcal{E}$ and N_ε a positive integer, a finite set of points $\{e_1, \ldots, e_{N_\varepsilon}\} \subset \mathcal{E}$ is called, for a given $\varepsilon > 0$, a ε-net of $\mathcal{S}_{\mathcal{E}}$ if :*

$$\mathcal{S}_{\mathcal{E}} \subseteq \cup_{j=1}^{N_\varepsilon} B(e_j, \varepsilon).$$

If $N_\varepsilon(\mathcal{S}_{\mathcal{E}})$ is the cardinality of the smallest ε-net (the minimal number of open balls of radius ε) in $\mathcal{E}$, needed to cover $\mathcal{S}_{\mathcal{X}}$, then we call Kolmogorov's entropy (metric entropy) of the set $\mathcal{S}_{\mathcal{E}}$, the quantity

$$\psi_{\mathcal{S}_{\mathcal{E}}}(\varepsilon) := \log N_\varepsilon(\mathcal{S}_{\mathcal{E}}).$$

From its name, one can figure that this concept of metric entropy was introduced by Kolmogorov [131] and was studied subsequently for numerous metric spaces. This concept was used by [80] to give sufficient conditions for continuity of Gaussian processes, and was the basis for striking generalizations of Donsker's theorem on the weak convergence of the empirical process. Suppose that $\mathcal{B}_{\mathcal{X}}$ and $\mathcal{S}_{\mathcal{X}}$ are two subsets of the semi-metric space $\mathcal{X}$ with Kolmogorov's entropy (for the radius ε) $\psi_{\mathcal{B}_{\mathcal{X}}}(\varepsilon)$ and $\psi_{\mathcal{S}_{\mathcal{X}}}(\varepsilon)$ respectively, then the Kolomogorov entropy for the subset $\mathcal{B}_{\mathcal{X}} \times \mathcal{S}_{\mathcal{X}}$ of the semi-metric space $\mathcal{X}^2$ by :

$$\psi_{\mathcal{B}_{\mathcal{X}}\times\mathcal{S}_{\mathcal{X}}}(\varepsilon) = \psi_{\mathcal{B}_{\mathcal{X}}}(\varepsilon) + \psi_{\mathcal{S}_{\mathcal{X}}}(\varepsilon).$$

Hence, $m\psi_{\mathcal{S}_{\mathcal{X}}}(\varepsilon)$ is the Kolmogorov entropy of the subset $\mathcal{S}_{\mathcal{X}}^m$ of the semi-metric space $\mathcal{X}^m$. Noting that if we designate by d the semi-metric on $\mathcal{X}$, then, we can define the semi metric on $\mathcal{X}^m$ by :

$$d_{\mathcal{X}^m}(\mathbf{x}, \mathbf{y}) := \frac{1}{m} d(x_1, y_1) + \cdots + \frac{1}{m} d(x_m, y_m)$$

for

$$\mathbf{x} = (x_1, \ldots, x_m),\ \mathbf{y} = (y_1, \ldots, y_m) \in \mathcal{X}^m.$$

Notice that the semi-metric plays an important role in this kind of study. The reader will find useful discussions about how to choose the semi-metric in [91] (see Chapters 3 and 13).

2.2. Conditions and comments

Let us present the conditions that we need in our analysis.

(C.1.) On the distributions/small-ball probabilities

(C.1.1) For $\mathbf{t} = (t_1, \ldots, t_m) \in \mathcal{X}^m$ and $\mathbf{h}(\mathbf{t}) = (h_1(t_1), \ldots, h_m(t_m)) \in \mathbb{R}_+^m \setminus \{0\}$, we have

$$\phi_{\mathbf{t}}(\mathbf{h}(\mathbf{t})) := \mathbb{P}(X_1 \in B(t_1, h_1(t_1)), \ldots, X_m \in B(t_m, h_m(t_m)))$$
$$0 < C_1 \tilde{\phi}(\mathbf{h}) f_1(\mathbf{t}) \leq \phi_{\mathbf{t}}(\mathbf{h}(\mathbf{t})) \leq C_2 \tilde{\phi}(\mathbf{h}) f_1(\mathbf{t}) < \infty,$$

where $f_1(\mathbf{t})$ is a non-negative functional in $\mathbf{t} = (t_1, \ldots, t_m) \in \mathcal{X}^m$, $\tilde{\phi}(\mathbf{h}) := \prod_{j=1}^m \phi(h_j(t_j))$ and $\phi(0) = 0$ and $\phi(u)$ is an invertible function absolutely continuous in a neighbor of the origin.

(C.1.2) For $i \neq j$, let $\mathbf{X_i} = (X_{i_1}, \ldots, X_{i_m})$, $\mathbf{X_j} = \left(X_{j_1}, \ldots, X_{j_m}\right)$ and $\mathbf{t} = (t_1, \ldots, t_m) \in \mathcal{X}^m$, we have

$$\sup_{\mathbf{i}\neq\mathbf{j}} \mathbb{P}\left\{\mathbf{X_i} \in \prod_{i=1}^m B(t_i, h_i(t_i)), \mathbf{X_j} \in \prod_{i=1}^m B(t_i, h_i(t_i))\right\} \leq \tilde{\Psi}(\mathbf{h}(\mathbf{t})) f_2(\mathbf{t}),$$

where $f_2(\mathbf{t})$ is a non-negative function, $\tilde{\Psi}(\mathbf{h}(\mathbf{t})) := \prod_{i=1}^m \psi(h_i(t_i))$ and $\tilde{\Psi}(\mathbf{h}(\mathbf{t})) \to 0$ as $n \to \infty$, satisfying $\tilde{\Psi}(\mathbf{h}(\mathbf{t})) \big/ \tilde{\phi}^2(\mathbf{h}(\mathbf{t}))$ is bounded.

(C.2) On the smoothness of the model

(C.2.1) The regression satisfies for $\varphi(\cdot) \in \mathscr{F}_m$, and for $1 \leq m \leq n$:

$$\exists \gamma > 0,\ \forall\, \mathbf{t}_1, \mathbf{t}_2 \in S_{\mathcal{X}}^m : \left| r^{(m)}(\varphi, \mathbf{t}_1) - r^{(m)}(\varphi, \mathbf{t}_2) \right| \leq C_3 d_{\mathcal{X}^m}^{\gamma}(\mathbf{t_1}, \mathbf{t_2}), \quad \text{for some} \quad \gamma > 0,$$

where for $1 \leq i \leq 2$: $\mathbf{t}_i := (t_{i_1}, \ldots, t_{i_m})$ and the semi metric $d_{\mathcal{X}^m}(\cdot, \cdot)$ on $\mathcal{X}^m$:

$$d_{\mathcal{X}^m}(\mathbf{x}, \mathbf{y}) := \frac{1}{m} d(x_1, \varphi(Y)) + \cdots + \frac{1}{m} d(x_m, y_m)$$

for $\mathbf{x} = (x_1, \ldots, x_m), \mathbf{y} = (y_1, \ldots, y_m) \in \mathcal{X}^m$.

(C.2.2) The conditional variance, defined for $u \in \mathcal{X}^m$, $\mathrm{Var}\left[\varphi(\mathbf{Y})|\mathbf{X}=\mathbf{u}\right] =: g_2(\mathbf{u},\varphi)$ is continuous in some neighborhood of $\mathbf{t}$

$$\sup_{d_{\mathcal{X}^m}(\mathbf{t},\mathbf{u})\leqslant h_n} |g_2(\mathbf{t},\varphi) - g_2(\mathbf{u},\varphi)| = o(1) \text{ as } n \to \infty.$$

Further, assume that for some $p > 2$, $\mathbb{E}|F(\mathbf{Y})|^p < \infty$, and $\varphi(\cdot) \in \mathscr{F}_m$,

$$g_p(\mathbf{t},\mathbf{u},\varphi) := \mathbb{E}\left(\left|\varphi(\mathbf{Y}) - r^{(m)}(\varphi,\mathbf{t})\right|^p \,\middle|\, \mathbf{X}=\mathbf{u}\right),$$

is continuous in some neighborhood of $\mathbf{t}$.

(C.2.3) For $\mathbf{u},\mathbf{v} \in \mathcal{X}^m$, the function $g_\varphi(\mathbf{t},\mathbf{u},\mathbf{v})$ does not depend on $\mathbf{i},\mathbf{j}$ and is continuous in some neighborhood of $(\mathbf{t},\mathbf{t})$

$$g_\varphi(\mathbf{t},\mathbf{u},\mathbf{v}) := \mathbb{E}\left(\left(\varphi(\mathbf{Y_i}) - r^{(m)}(\varphi,\mathbf{t})\right)\left(\varphi\left(\mathbf{Y_j}\right) - r^{(m)}(\varphi,\mathbf{t})\right) \,\middle|\, \mathbf{X_i}=\mathbf{u},\mathbf{X_j}=\mathbf{v}\right).$$

(C.3) On the kernel function

(C.3.1) The kernel functions $K(\cdot)$ is supported within $[0,1]$ and there exists some constants $0 < \kappa_1 \leq \kappa_2 < \infty$, such that

$$\int_0^1 K(x)dx = 1,$$

and

$$0 < \kappa_1 \mathbb{1}_{[0,1]}(\cdot) \leq K(\cdot) \leq \kappa_2 \mathbb{1}_{[0,1]}(\cdot).$$

(C.3.2) The kernel $K(\cdot)$ is a positive function and differentiable function on $[0,1]$ with derivative $K'(\cdot)$ such that

$$-\infty < \kappa_3 < K'(\cdot) < \kappa_4 < 0. \tag{2.8}$$

(C.4) On the classes of functions

(C.4.1) The class of functions $\mathscr{F}_m$ is bounded and its envelope function satisfies for some $0 < M < \infty$:

$$F(\mathbf{y}) \leq M, \qquad \mathbf{y} \in \mathcal{Y}^m$$

(C.4.2) The class of functions $\mathscr{F}_m$ is unbounded and its envelope function satisfies for some $p > 2$:

$$\theta_p := \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \mathbb{E}\left(F^p(\mathbf{Y})|\mathbf{X}=\mathbf{t}\right) < \infty.$$

(C.4.3) The metric entropy of the class $\mathscr{F}_m\mathscr{K}^m$ satisfies, for some $1 \leq p < \infty$:

$$\int_0^\infty (\log N(u, \mathscr{F}_m\mathscr{K}^m, \|\cdot\|_p))^{\frac{1}{2}} du < \infty,$$

where

$$\mathscr{F}_m\mathscr{K}^m = \{fg : f \in \mathscr{F}_m, g \in \mathscr{K}^m\}.$$

(C.4.4) The class of functions $\mathscr{F}_m\mathscr{K}^m$ is supposed to be of VC-type with envelope function previously defined. Hence, there are two finite constants b and ν such that:

$$N\left(\epsilon, \mathscr{F}_m\mathscr{K}^m, \|\cdot\|_{L_2(Q)}\right) \leq \left(\frac{b\|F\kappa^m\|_{L_2(Q)}}{\epsilon}\right)^{\nu}$$

for any $\epsilon > 0$ and each probability measure such that $Q(F)^2 < \infty$.

(C.5) On the dependence of the random variables

(C.5.1) Absolute regularity

$$\sum_{s=1}^{\infty} s^{\delta}\left(\log(s)\right)^{\delta(1-1/p)}\left(\beta(s)\right)^{1-2/p} < \infty,$$

for some $p > 2$ and $\delta > 1 - 2/p$.

(C.5.2) There is a sequence of positive integers $\{s_n\}_{n\in\mathbb{N}^*}$ such that, as $n \to \infty$,

$$s_n \to \infty, \qquad s_n = o\left(\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\right), \qquad \left(\frac{n}{\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\right)^{1/2}\beta\left(s_n\right) \longrightarrow 0.$$

(C.6.) On the entropy

For n large enough and for some $\omega > 1$, the Kolomogorov's entropy satisfies :

$$\frac{(\log n)^2}{n\phi(h_K)} \leq m\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right) \leq \frac{n}{\log n}\phi(h_K), \tag{2.9}$$

$$\sum_{n=1}^{\infty} \exp\left\{m(1-\omega)\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)\right\} < \infty, \tag{2.10}$$

where $\psi_{S_{\mathcal{X}}}(\varepsilon) := \log N_{\varepsilon}(S_{\mathcal{X}})$, and $N_{\varepsilon}(S_{\mathcal{X}})$ is the minimal number of open balls of radius ε in $\mathcal{X}$, needed to cover $S_{\mathcal{X}}$.

(C.7.) The sequences $\{\widetilde{h}_n\}$ and $\{\widetilde{h'}_n\}$ (resp. $\{h_{n,1}\}$ and $\{h_{n,2}\}$) verify

$$\widetilde{h'}_n \longrightarrow 0 \text{ and } \frac{(\log n/n)}{\min\{\widetilde{h}_n^2, \phi^2(\widetilde{h}_n)\}} \longrightarrow 0 \text{ as } n \longrightarrow \infty. \tag{2.11}$$

(C.8.) There exist sequences $\{\rho_{n,1}, \ldots, \rho_{n,m}\} \subset (0,1)^m$, $\{k_{1,n}\} \subset \mathbb{Z}^+$ and $\{k_{2,n}\} \subset \mathbb{Z}^+$ ($k_{1,n} \leq k \leq k_{2,n}$) and constants $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_m)$, such that

$$0 < \mu_j \leq \nu_j < \infty, \text{ for all } j = 1, \ldots, m \text{ (and we note } \boldsymbol{\mu} \leq \boldsymbol{\nu}),$$

and

$$\mu_j\phi^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right) \leq \phi_{t_j}^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right) \text{ and } \phi_{t_j}^{-1}\left(\frac{k_{2,n}}{\rho_{n,j}n}\right) \leq \nu_j\phi^{-1}\left(\frac{k_{2,n}}{\rho_{n,j}n}\right), \tag{2.12}$$

$$\phi^{-1}\left(\frac{k_{2,n}}{\rho_{n,j}n}\right) \longrightarrow 0, \tag{2.13}$$

$$\min\left\{\frac{1-\rho_{n,j}}{4}\frac{k_{1,n}}{\ln n}, \frac{(1-\rho_{n,j})^2}{4\rho_{n,j}}\frac{k_{1,n}}{\ln n}\right\} > 2, \tag{2.14}$$

$$\frac{(\log n/n)}{\min\left\{\mu_j\phi^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right), \phi\left(\mu_j\phi^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right)\right)\right\}} \longrightarrow 0. \tag{2.15}$$

Additional/alternative conditions

(C.1'.) For $\mathbf{X_i} = (X_{i_1}, \ldots, X_{i_m})$, $\mathbf{X_j} = \left(X_{j_1}, \ldots, X_{j_m}\right)$ and $\mathbf{t} = (t_1, \ldots, t_m) \in \mathcal{X}^m$:

(C.1'.1) We have:

$$\phi_{\mathbf{t}}(\mathbf{x}) = \tilde{\phi}(\mathbf{x}) f_1(\mathbf{t}) < \infty \text{ as } \mathbf{x} \to \infty.$$

with $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}_+^m$, $\phi(0) = 0$, and $\phi(u)$ is absolutely continuous in a neighborhood of the origin.

(C.1'.2) we have:

$$\sup_{\mathbf{i}\neq\mathbf{j}} \mathbb{P}\left\{\mathbf{X_i} \in \prod_{i=1}^m B(t_i, h_i(t_i)), \mathbf{X_j} \in \prod_{i=1}^m B(t_i, h_i(t_i))\right\} \leq \tilde{\Psi}(\mathbf{h}(\mathbf{t})) f_2(\mathbf{t}),$$

where $f_2(\mathbf{t})$ is a non-negative function, $\tilde{\Psi}(\mathbf{h}(\mathbf{t})) := \prod_{i=1}^m \psi(h_i(t_i))$, and $\tilde{\Psi}(\mathbf{h}(\mathbf{t})) \to 0$ as $n \to \infty$, satisfying $\tilde{\Psi}(\mathbf{h}(\mathbf{t})) \Big/ \tilde{\phi}^2(\mathbf{h}(\mathbf{t}))$ is bounded.

(C.3'.) The kernel functions $K(\cdot)$ is supported within $[0, 1]$ and there exists some constants $0 < \kappa_2$, $0 < \kappa_1' \leq \kappa_2' < \infty$, such that for $j = 1, 2$:

$$\int_0^1 K(x)dx = 1, \quad K(\cdot) \leq \kappa_2 \mathbb{1}_{[0,1]}(\cdot), \quad \frac{h_n}{\phi(h_K(t))}\int_0^1 K^j(x)\phi'(vh_K)dv \to \kappa_j' \text{ as } n \to \infty.$$

Comments

In our nonparametric functional regression model, we deal with a complex theoretical challenge. Mainly establishing functional central limit theorems for the conditional empirical processes and the conditional U-process under functional absolute regular data. We also use random (or data-dependent) bandwidths based on the k nearest neighbors (kNN) approaches. Although standard statistical methods cannot be utilized in the functional setting, most imposed conditions overlap with some characteristics of the infinite-dimensional spaces, such as the topological structure of $\mathcal{X}^m$, the probability distribution of X, and the measurability concept for the classes $\mathcal{F}_m$ and $\mathcal{K}^m$. It is worth mentioning that most of the conditions that we will be using throughout this paper are inspired by [43, 91, 99, 141, 147]. Let us start with assumption **(C.1.1)**, which was adapted from [147], who in turn was inspired by [99]. As explained by [147], if $\mathcal{X}^m = \mathbb{R}^m$, then condition **(C.1.1)** coincides with the fundamental axioms of probability calculus. Furthermore, if $\mathcal{X}^m$ is an infinite-dimensional Hilbert space, then $\phi(h_K)$ can converge to 0 exponentially as $n \to \infty$. Condition **(C.1.1)** can be considered a standard condition on the small ball probability, which is used to control the behavior of $\phi_{\mathbf{t}}(\cdot)$ around zero. It shows that we can approximately write the small ball probability as a product of two independent functions $\tilde{\phi}(\cdot)$ and $f_1(\cdot)$; see, for instance, [148] for the diffusion process, [18] for a Gaussian measure, and [139] for a general Gaussian process. The most frequent result available in the literature is of the form $\varphi_t(\varepsilon) \sim g(t)\phi(\varepsilon)$ where $\phi(\varepsilon) = \varepsilon^{\gamma}\exp(-C/\varepsilon^{p})$ with $\gamma \geq 0$ and $p \geq 0$. It corresponds to the Ornstein-Uhlenbeck and general diffusion processes (for such processes, $p = 2$ and $\gamma = 0$) and the fractal processes (for such processes, $\gamma > 0$ and $p = 0$). For more examples, refer to [92]. It is worth noting that, in general, when we deal with functional data, we need some information about the variability of the small-ball probability to adapt to the bias of nonparametric estimators; this information is usually obtained by supposing that:

(C.1.1")

$$\forall u \in [0,1]: \qquad \lim_{r\to\infty}\frac{\phi_t(ur)}{\phi_t(r)} = \lim_{r\to\infty}\mathbb{P}\Big(d(X,t) \leqslant ur \Big| d(X,t) \leqslant r\Big) =: \tau_t(u) < \infty.$$

Condition **(C.2)** concerns the regularity of the model; it consists of mild conditions on the continuity of certain conditional moments, in addition to the standard Lipschitz assumption on the regression ((**C.2.1)**. Assumption **(C.3)** is another classical condition in nonparametric estimation models which concerns the kernel function $K(\cdot)$. It should be noted that condition **(C.3.1)** can be replaced with condition **(C.3')** to find an expression for the asymptotic variance. We use condition **(C.4.1)** when dealing with bounded functions. However, our interest also extends to conditional U-processes indexed by an unbounded class of functions. In this case, we replace **(C.4.1)** by **(C.4.2)**. Keep in mind that there is a trade-off between the moment order p in **(C.4.2)** and the decay rate of the mixing coefficient $\beta(s)$ imposed in **(C.5)**: the larger p is, the weaker the decay of $\beta(s)$. Also note that if $\beta(s) = \mathrm{e}^{-as}$, i.e, $\beta(s)$ decays exponentially fast, then **(C.5.)** is automatically satisfied. Furthermore, this condition is indispensable in our work since studying the weak convergence of the empirical processes entails establishing asymptotic equi-continuity. For Assumption **(C.4.4)**, see [163, Examples 26 and 38], [157, Lemma 22], [82, §4.7.], [193, Theorem 2.6.7], [132, §9.1] provide a number of sufficient conditions under which **(C.4.4)** holds, we may also refer to [70, §3.2] for further discussions. For instance, it is satisfied, for general $d \geq 1$, whenever $g(\mathbf{x}) = \phi(p(\mathbf{x}))$, with $p(\mathbf{x})$ being a polynomial in d variables and $\phi(\cdot)$ being a real-valued function of bounded variation, we refer the reader to [88, p. 1381]. We also mention that the class of function is assumed to be in general pointwise measurable, that is satisfied whenever $K(\cdot)$ is of bounded variation on $\mathbb{R}^d$ (in the sense of Hardy and Kauser [110, 133, 195], see, e.g., [61, 114, 156, 196]). The condition **(C.6.)** takes into account the topological considerations by controlling the Kolmogorov entropy[§] of the set $S_{\mathcal{X}}^{m}$, which is standard in nonparametric models when we study the uniform consistency and the uniform in bandwidth consistency, we refer to [93] and [134] for discussions. As mentioned in [134], there are special cases of functional spaces $\mathcal{X}$ and subsets $S_{\mathcal{X}}$ where $\psi_{S_{\mathcal{X}}}(\log(n)/n) \gg \log(n)$. Some examples are the closed ball in a Sobolev space, the unit ball of the Cameron-Martin space, and a compact subset in a Hilbert space with a projection semi-metric (see [93, 131, 192], respectively, for further details). In all these cases it is easy to see that **(C.6.)** is verified

[§]The concept of metric entropy was introduced by Kolmogorov (*cf.* [131]) and studied subsequently for numerous metric spaces. This concept was used by [80] to give sufficient conditions for the continuity of Gaussian processes, and was the basis for striking generalizations of Donsker's theorem on the weak convergence of the empirical process, refer to [191] for the connection of this notion with Le Cam's work.

as soon as $\beta > 2$. Assumption **(C.7.)** is essential to establish the rates of convergence (consistency) of the estimator defined in (2.5), while assumption **(C.8.)** adapts condition **(C.7.)** to the case of the functional conditional U-statistics in the k-NN setting.

Remark 2.4. *Note that the condition **(C.4.2)** may be replaced by more general hypotheses upon moments of* $\mathbf{Y}$ *as in [70]. That is*

(M.1)$''$ We denote by $\{\mathcal{M}(x) : x \geq 0\}$ a nonnegative continuous function, increasing on $[0, \infty)$, and such that, for some $s > 2$, ultimately as $x \uparrow \infty$,

$$(i)\ x^{-s}\mathcal{M}(x) \downarrow; (ii)\ \ x^{-1}\mathcal{M}(x) \uparrow . \tag{2.16}$$

For each $t \geq \mathcal{M}(0)$, we define $\mathcal{M}^{inv}(t) \geq 0$ by $\mathcal{M}(\mathcal{M}^{inv}(t)) = t$. We assume further that:

$$\mathbb{E}\left(\mathcal{M}\left(|F(\mathbf{Y})|\right)\right) < \infty.$$

The following choices of $\mathcal{M}(\cdot)$ are of particular interest:

(i) $\mathcal{M}(x) = x^p$ for some $p > 2$;
(ii) $\mathcal{M}(x) = \exp(sx)$ for some $s > 0$.

3. Uniform consistency

For simplicity reasons, the condition **(C.1.1)** on the small ball probability will be replaced by:

(H.1) For $\mathbf{h} \in \mathbb{R}_+^m \backslash \{0\}$ and $\mathbf{t} \in \mathcal{X}^m$

$$0 < C_1\tilde{\phi}(\mathbf{h}) \leq \phi_{\mathbf{t}}(\mathbf{h}(\mathbf{t})) \leq C_2\tilde{\phi}(\mathbf{h}) < \infty, \tag{3.1}$$

this standard condition can be considered an extension of the multivariate case where we assume that the density function of the variable $\mathbf{X}$ is strictly positive. Also, it is worth mentioning that, if in particular, we denote $\mathbf{h}_K := (h_K, \ldots, h_K) \in \widetilde{(h_{n,1}, h_{n,2})}^m$, we can find two positive constants C_1', C_2', such that

$$0 < C_1'\phi(h_K) \leq \phi_{\mathbf{t}}(\mathbf{h}_K) \leq C_2'\phi(h_K) < \infty, \tag{3.2}$$

which is similar to condition **(C.1)** used in [42, 181], so we will deal with $\phi(h_K)$ instead of $\phi^m(h_K)$, (whenever we encounter a similar situation in the proofs). This approach is not only for notational purposes but also to make it easier to bridge the UIB and the UINN results.

3.1. Uniform consistency of the kNN kernel estimator for regression

In this section, we consider the uniform consistency of the functional regression operator in its general form, which is given for all $t \in \mathcal{X}$, by

$$\widehat{r}_n^{*(1)}(\varphi, \mathbf{t}, H_{n,k}(t)) = \frac{\sum_{i=1}^{n} \varphi(Y_i) K\left(\frac{d(t, X_i)}{H_{n,k}(t)}\right)}{\sum_{i=1}^{n} K\left(\frac{d(t, X_i)}{H_{n,k}(t)}\right)}, \tag{3.3}$$

where $k = k_n$ depending on n and

$$H_{n,k}(t) = \min\left\{h \in \mathbb{R}^+ \text{such that} \sum_{i=1}^{n} \mathbb{1}_{B(t,h)}(X_i) = k\right\}.$$

In fact, the k-NN operator presented in (3.3) can be considered as a generalization of the usual kernel regression

$$\widehat{r}_n^{(1)}(\varphi, t, h_K) = \frac{\sum_{i=1}^{n} \varphi(Y_i) K\left(\frac{d(t, X_i)}{h_K}\right)}{\sum_{i=1}^{n} K\left(\frac{d(t, X_i)}{h_K}\right)}, \quad \text{for each } \ t \in \mathcal{X}, \tag{3.4}$$

where the bandwidth $h_K \in \mathbb{R}_+^*$ depends on n (but does not depend on t).

3.1.1. UIB consistency for functional regression

Recall the bandwidths $h_{n,1}$ and $h_{n,2}$ given in the condition **(C.7.)**. The following theorem will play an instrumental role in the sequel.

Theorem 3.1. *Under the assumptions* **(H.1.)**,**(C.2.1)**, **(C.3.1)**,**(C.4.1)**,**(C.4.3)**, **(C.6.)** *and* **(C.7.)** *(for* $m = 1$*), we have, as* $n \to \infty$,

$$\sup_{\varphi K \in \mathscr{F}\mathscr{K}} \sup_{h_{n,1} \leq h_K \leq h_{n,2}} \sup_{t \in S_{\mathcal{X}}} \left|\widehat{r}_n^{(1)}(\varphi, t; h_K) - r^{(1)}(\varphi, t)\right| = O\left(h_{n,2}^{\gamma}\right) + O_{a.co}\text{¶}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right). \tag{3.5}$$

The following result gives uniform consistency when the class of functions is unbounded.

Corollary 3.2. *Under the assumptions* **(H.1.)**, **(C.2.1)**, **(C.3.1)**,**(C.4.2)**, **(C.4.3)**, **(C.6.)** *and* **(C.7.)** *(for* $m = 1$*), we have*

$$\sup_{\varphi K \in \mathscr{F}\mathscr{K}} \sup_{h_{n,1} \leq h_K \leq h_{n,2}} \sup_{t \in S_{\mathcal{X}}} \left|\widehat{r}_n^{(1)}(\varphi, t; h_K) - r^{(1)}(\varphi, t)\right| = O\left(h_{n,2}^{\gamma}\right) + O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right). \tag{3.6}$$

¶Let (z_n) for $n \in \mathbb{N}$, be a sequence of real r.v.'s. We say that (z_n) converges almost-completely (a.co.) toward zero if, and only if, for all

$$\epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}\left(|z_n| > \epsilon\right) < \infty.$$

Moreover, we say that the rate of the almost-complete convergence of (z_n) toward zero is of order u_n (with $u_n \to 0$) and we write $z_n = O_{a.co.}(u_n)$ if, and only if, there exists $\epsilon > 0$ such that

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|z_n| > \epsilon u_n\right) < \infty.$$

This kind of convergence implies both the almost-sure convergence and the convergence in probability.

3.1.2. UINN consistency for functional regression

Now, we can state the main results of this section concerning the k-NN functional regression. Recall the bandwidths $k_{1,n}$ and $k_{2,n}$ given in the condition **(C.8.)**.

Theorem 3.3. *Under the assumptions **(H.1.)**, **(C.2.1)**, **(C.3.1)**,**(C.4.1)**, **(C.4.3)**, **(C.6.)** and **(C.7.)** (for $m = 1$), if, in addition, condition **(C.8.)** is satisfied, then, we have*

$$\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{k_{1,n}\le k\le k_{2,n}}\ \sup_{t\in S_X}\left|\widehat{r}_n^{*(1)}(\varphi,t;h_{n,k}(t)) - r^{(1)}(\varphi,t)\right| = O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n n}\right)^{\gamma}\right) + O_{a.co}\left(\sqrt{\frac{\psi_{S_X}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right).$$

The following result gives uniform consistency when the class of functions is unbounded.

Corollary 3.4. *Under the assumptions **(H.1.)**, **(C.2.1)**, **(C.3.1)**,**(C.4.2)**, **(C.4.3)**, **(C.6.)** and **(C.7.)** (for $m = 1$), and if condition **(C.8.)** is satisfied, then we have*

$$\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{k_{1,n}\le k\le k_{2,n}}\ \sup_{t\in S_X}\left|\widehat{r}_n^{*(1)}(\varphi,t;h_{n,k}(t)) - r^{(1)}(\varphi,t)\right| = O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n n}\right)^{\gamma}\right) + O_{a.co}\left(\sqrt{\frac{\psi_{S_X}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right). \tag{3.7}$$

3.2. Relative-error prediction

Recall that the operator m is usually estimated by minimizing the expected squared loss function $\mathbb{E}\left[(Y - m(X))^2 \mid X\right]$. Nonetheless, this loss function, which is regarded as a measure of prediction performance, may be inappropriate in certain circumstances. In fact, the application of least-squares regression translates to giving all variables in the study equal weight. Consequently, the prevalence of outliers can render results irrelevant. In this paper, we, therefore, circumvent the limitations of classical regression by estimating the operator m with regard to the minimization of the mean squared relative error (MSRE):

$$\mathbb{E}\left[((Y - m(X))/Y)^2 \mid X\right] \quad \text{for } Y > 0 \ \ a.s. \tag{3.8}$$

This criterion is clearly a more meaningful measure of the prediction performance than the least square error, in particular, when the range of predicted values is large. Moreover, the solution of (3.8) can be explicitly expressed by the ratio of the first two conditional inverse moments of Y given X. In fact, in order to construct the regression estimator allowing the best MSRE prediction, we assume that the first two conditional inverse moments of Y given X, that is $g_\gamma(x) := \mathbb{E}\left(Y^{-\gamma} \mid X = x\right)$ for $\gamma = 1, 2$, exist and are finite almost-surely (a.s.). Then, one can show easily, cf. [74, 123, 159], that the best mean squared relative error predictor of Y given X is:

$$\breve{r}(t) = \mathbb{E}\left(Y^{-1} \mid X = t\right)/\mathbb{E}\left(Y^{-2} \mid X = t\right) = g_1(t)/g_2(t), \quad \text{a.s.}$$

Thus, we estimate the regression operator $\breve{r}(\cdot)$, which minimizes the MSRE by:

$$\breve{r}_n^{(1)}(t, h_K) = \frac{\sum_{i=1}^{n} Y_i^{-1} K\left(\frac{d(t, X_i)}{h_K}\right)}{\sum_{i=1}^{n} Y_i^{-2} K\left(\frac{d(t, X_i)}{h_K}\right)}, \quad \text{for each } \ t \in \mathcal{X}, \tag{3.9}$$

and

$$\breve{r}_n^{*(1)}(t, h_K) = \frac{\sum_{i=1}^{n} Y_i^{-1} K\left(\frac{d(t, X_i)}{H_{n,k}(t)}\right)}{\sum_{i=1}^{n} Y_i^{-2} K\left(\frac{d(t, X_i)}{H_{n,k}(t)}\right)}, \quad \text{for each } \ t \in \mathcal{X}. \tag{3.10}$$

By considering the special cases $\varphi(y) = y^{-1}$ and $\varphi(y) = y^{-2}$ in Corollaries 3.2 and 3.4, we obtain the following results complementing the work of [22, 23, 24, 74].

Corollary 3.5. *Under the assumptions* ***(H.1.)****,* ***(C.2.1)****,* ***(C.3.1)****,****(C.4.2)****,* ***(C.4.3)****,* ***(C.6.)*** *and* ***(C.7.)*** *(for* $m = 1$*), we have*

$$\sup_{K\in\mathscr{K}} \sup_{h_{n,1}\le h_K\le h_{n,2}} \sup_{t\in S_{\mathcal{X}}} \left|\breve{r}_n^{(1)}(t, h_K) - \breve{r}(t)\right| = O\left(h_{n,2}^{\gamma}\right) + O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right). \tag{3.11}$$

The following result is not considered in the literature.

Corollary 3.6. *Under the assumptions* ***(H.1.)****,* ***(C.2.1)****,* ***(C.3.1)****,****(C.4.2)****,* ***(C.4.3)****,* ***(C.6.)*** *and* ***(C.7.)*** *(for* $m = 1$*), and if condition* ***(C.8.)*** *is satisfied, then we have*

$$\begin{aligned}\sup_{K\in\mathscr{K}} \sup_{k_{1,n}\le k\le k_{2,n}} \sup_{t\in S_{\mathcal{X}}} \left|\breve{r}_n^{*(1)}(t, h_K) - \breve{r}(t)\right| \quad &= \quad O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n n}\right)^{\gamma}\right) \\ &\quad + O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right).\end{aligned} \tag{3.12}$$

3.3. Uniform consistency of the kNN functional conditional U-statistics

In addition to the conditions imposed before, the following assumptions are essential to obtain exponential inequalities for dependent data later in the proofs:

(A1) Assume $\{X_i\}_{i\in\mathbb{N}^*}$ is strictly stationary, and there exists an absolute constant $\delta \ge 1$ such that for any $n \ge 1$, we have the β-mixing coefficient, corresponding to $\{X_i\}_{i\in\mathbb{N}^*}$, satisfies $\beta(n) \lesssim n^{-\delta}$.

(A2) Assume, uniformly, for any integer J such that $1 \leq J \leq m-1$ and arbitrary $1 \leq i_1 < \cdots < i_J \leq n$, conditional on $X_{i_1}, \ldots, X_{i_J}$, the sequence $\{X_i\}_{i=i_J+1}^{\infty}$ satisfies, for the α-mixing coefficient corresponding to it,

$$\alpha\left(n; X_{i_1}, \ldots, X_{i_J}\right) := \sup_{j \geq i_J+1} \alpha\left(\sigma_{i_J+1}^{j}, \sigma_{j+n}^{\infty}; \mathbb{P}\left(\cdot \mid X_{i_1}, \ldots, X_{i_J}\right)\right) \lesssim \exp(-\gamma n), \text{ a.s.}$$

where $\mathbb{P}\left(\cdot \mid X_{i_1}, \ldots, X_{i_J}\right)$ stands for the conditional probability. In particular, we have, for the α-mixing coefficient corresponding to $\{X_i\}_{i \in \mathbb{N}^*}$ itself

$$\alpha(n) \lesssim \exp(-\gamma n).$$

The β-mixing condition **(A1)** is typically necessary to obtain asymptotic normality for U-statistics in the absence of a strict Lipchitz-continuity assumption for the kernel functions, for instance, see [72, 203] and Remarks 2.2 and 2.3 of [107]. Assumption **(A2)** is often less restrictive than the ϕ-mixing condition. As will be shown in [107], finite-state and vector-valued absolutely continuous data sequences of exponentially ϕ-mixing decaying rate satisfy **(A2)**. We note that this is more restrictive than the polynomially mixing decaying rate. [203]. This is because we need to calculate higher moments of the U-statistics to obtain sharp concentration inequality. The "exponentially decaying rate" condition is routine in the literature of deriving concentration inequalities for weakly dependent data, see [149] and Remarks 2.4 and 2.5 in [107].

This section considers the uniform consistency, the UIB, and the UINN consistency of the functional conditional U-statistic given by (2.1). First, let's introduce some notation. For some interval $\mathcal{H}_n^{(m)} \subset \mathbb{R}_+^m \setminus \{0\}$, we denote

$$\mathcal{H}_n^{(m)} := \prod_{j=1}^{m} (h_{n,j}, h'_{n,j}),$$

where

$$0 < h_{n,j} < h'_{n,j} \text{ and } \lim_{n\to\infty} h_{n,j} = \lim_{n\to\infty} h'_{n,j} = 0 \quad , \forall j = 1, \ldots, m.$$

In the sequel, we denote (unless stated otherwise)

$$\widetilde{h}_n = \min_{1 \leq j \leq m} h_{n,j} \text{ and } \widetilde{h}'_n = \max_{1 \leq j \leq m} h'_{n,j}.$$

For all $\mathbf{b} = (b_1, \ldots, b_m) \in (0,1)^m$, let us denote

$$\mathcal{H}_0^{(m)} := \prod_{j=1}^{m} \left(h_{n,j}, b_j\right),$$

and

$$\widetilde{b}_0 := \max_{1 \leq j \leq m} b_j.$$

For notational convenience, in the case of $m = 1$, we denote $\mathcal{H}_n^{(1)}$ and $\mathcal{H}_0^{(1)}$ simply $\mathcal{H}_n$ and $\mathcal{H}_0$. The same goes for other similar notation unless stated otherwise. Set

$$\mathbf{X} \ := \ (X_1, \ldots, X_m) \in \mathcal{X}^m, \quad \mathbf{Y} := (Y_1, \ldots, Y_m) \in \mathcal{Y}^m,$$

$$
\begin{aligned}
\mathbf{X}_i &:= (X_{i_1},\ldots,X_{i_m}), \quad \mathbf{Y}_i := (Y_{i_1},\ldots,Y_{i_m}),\\
\mathbf{h}_{n,k}(\mathbf{t}) &:= (H_{n,k}(t_1),\ldots,H_{n,k}(t_m)) \quad \text{for} \quad \mathbf{t}=(t_1,\ldots,t_m)\in S^m_{\mathcal{X}},\\
G_{\varphi,\mathbf{t},\mathbf{h}}(\mathbf{x},\mathbf{y}) &:= \frac{\varphi(y_1,\ldots,y_m)\prod_{i=1}^m K\left(\frac{d(x_i,t_i)}{h_i(t_i)}\right)}{\prod_{i=1}^m \mathbb{E}\left[K\left(\frac{d(X_i,t_i)}{h_i(t_i)}\right)\right]} \quad \text{for } \mathbf{t},\mathbf{x}\in\mathcal{X}^m, \mathbf{y}\in\mathcal{Y}^m;\\
\mathcal{G} &:= \left\{G_{\varphi,\mathbf{t},\mathbf{h}}(\cdot,\cdot)\ \varphi\in\mathscr{F}_m,\ \mathbf{t}=(t_1,\ldots,t_m)\in\mathcal{X}^m\right\},\\
\mathcal{G}^{(p)} &:= \left\{\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}(\cdot,\cdot),\ \varphi\in\mathscr{F}_m,\ \mathbf{t}=(t_1,\ldots,t_m)\right\},\\
u_n(\varphi,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t})) &= u_n^{(m)}(G_{\varphi,\mathbf{t},\mathbf{h}_{n,k}}) := \frac{(n-m)!}{n!}\sum_{i\in I(m,n)} G_{\varphi,\mathbf{t},\mathbf{h}_{n,k}}(\mathbf{X}_i,\mathbf{Y}_i),
\end{aligned}
$$

and for some symmetric measurable function $f(\cdot)$ define the $\mathbb{P}$-canonical function $\pi_{p,m}$, $p=1,\ldots,m$, see [10] and [68] (we replace the index k with p to avoid confusing it with the smoothing parameter k), by

$$\pi_{p,m}f(t_1,\ldots,t_p) := (\delta_{t_1}-\mathbb{P})\cdots(\delta_{t_p}-\mathbb{P})\mathbb{P}^{m-p}f,$$

where for measures Q_i on S we let

$$Q_1\cdots Q_m h = \int_{S^m} h(x_1,\ldots,x_m)dQ_1(x_1)\cdots dQ_m(x_m),$$

and δ_x denote Dirac measure at point $x\in\mathcal{X}$. This decomposition follows easily by expanding

$$f(\mathbf{x}_1,\ldots,\mathbf{x}_m) = \delta_{\mathbf{x}_1}\times\cdots\times\delta_{\mathbf{x}_m}f = ((\delta_{\mathbf{x}_1}-\mathbb{P})+\mathbb{P})\times\cdots\times((\delta_{\mathbf{x}_m}-\mathbb{P})+\mathbb{P})f,$$

into terms of the form

$$\left(\delta_{\mathbf{x}_{i_1}}-\mathbb{P}\right)\times\cdots\times\left(\delta_{\mathbf{x}_{i_m}}-\mathbb{P}\right)\times\mathbb{P}^{m-\ell}f.$$

It is very simple to check that $f(\cdot)$ symmetric is $\mathbb{P}$-degenerate of order $r-1$ ‖ if $r=\min\{\ell>0:\pi_\ell f\not\equiv 0\}$. For example,

$$\pi_{1,m}h(x) = \mathbb{E}(h(X_1,\ldots,X_m)\mid X_1=x) - \mathbb{E}h(X_1,\ldots,X_m).$$

‖

Definition 3.7. *A $\mathbb{P}^m$-integrable symmetric function of m variables, $f: S^m\mapsto\mathbb{R}$, is $\mathbb{P}$-degenerate of order $r-1, 1<r\leq m$, if $\int f(\mathbf{x}_1,\ldots,\mathbf{x}_m)\,d\mathbb{P}^{m-r+1}(\mathbf{x}_r,\ldots,\mathbf{x}_m) = \int f d\mathbb{P}^m$ for all $\mathbf{x}_1,\ldots,\mathbf{x}_{r-1}\in S$ whereas*

$$\int f(\mathbf{x}_1,\ldots,\mathbf{x}_m)\,d\mathbb{P}^{m-r}(\mathbf{x}_{r+1},\ldots,\mathbf{x}_m)$$

is not a constant function. If f is $\mathbb{P}^m$-centered and is P-degenerate of order $m-1$, that is, if

$$\int f(\mathbf{x}_1,\ldots,\mathbf{x}_m)\,d\mathbb{P}(\mathbf{x}_1)=0 \quad \text{for all } \mathbf{x}_2,\ldots,\mathbf{x}_m\in S,$$

then $f(\cdot)$ is said to be canonical or completely degenerate with respect to $\mathbb{P}$. If $f(\cdot)$ is not degenerate of any positive order, we say it is nondegenerate or degenerate of order zero.

It's clear that, for all $\varphi \in \mathscr{F}_m$:

$$\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t})) = \frac{u_n(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t}))}{u_n(1, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t}))},$$

and $u_n(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t}))$ is a classical U-statistic with the U- kernel $G_{\varphi,\mathbf{t},\mathbf{h}_{n,k}}(\mathbf{x}, \mathbf{y})$. However, the study of the uniform consistency of $\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t}))$ to $r^{(m)}(\varphi, \mathbf{t})$ can not be done with a straightforward approach due to the randomness of the bandwidth vector $\mathbf{h}_{n,k}(\mathbf{t})$ which poses some technical problems. To circumvent this, our strategy is first to study the uniform consistency of $\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h})$, where $\mathbf{h} = (h_1, \ldots, h_m) \in \mathcal{H}_n^{(m)}$ is a multivariate bandwidth that does not depend on $\mathbf{t}$ and k. Hence, we study the uniform consistency and the UIB consistency of $u_n(\varphi, \mathbf{t}, \mathbf{h})$ to $\mathbb{E}(u_n(\varphi, \mathbf{t}, \mathbf{h}))$ when $\varphi \in \mathscr{F}_m$ and when $\varphi \equiv 1$, and we shall consider an appropriate centering factor than the expectation $\mathbb{E}\left(\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h})\right)$, hence we define :

$$\widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h})\right) = \frac{\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)}{\mathbb{E}\left(u_n(1, \mathbf{t}, \mathbf{h})\right)}. \tag{3.13}$$

The second step will be the use a general lemma given in [44], adapted to our setting, similar to that of [134] (see Subsection 8.1.1) to derive the results for the bandwidth $\mathbf{h}_{n,k}(\mathbf{t})$.

3.3.1. Uniform consistency and UIB consistency for a multivariate bandwidth

Next, we will give the UIB results for all $\mathbf{t} \in S_{\mathcal{X}}^m$ and $\mathbf{h} \in \mathcal{H}_n^{(m)}$. We first start with announcing the result concerning the uniform derivation of the estimate $u_n(\varphi, \mathbf{t}, \mathbf{h})$ with respect to $\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)$ when the class of functions is bounded.

Theorem 3.8. *Suppose that the conditions* ***(H.1.), (C.3.1), (C.4.1), (C.4.4), (C.6.)*** *and* ***(C.7.)*** *are fulfilled, we infer that, as $n \to \infty$,*

$$\sup_{\varphi\widetilde{K} \in \mathscr{F}_m \mathscr{K}^m} \sup_{\mathbf{h} \in \mathcal{H}_n^{(m)}} \sup_{\mathbf{t} \in S_{\mathcal{X}}^m} \left|u_n(\varphi, \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right). \tag{3.14}$$

The following result covers the uniform derivation of the estimate $u_n(\varphi, \mathbf{t}, \mathbf{h})$ with respect to $\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)$ when the class of functions is unbounded satisfying general moments condition.

Theorem 3.9. *Suppose that the conditions* ***(H.1.), (C.3.1), (C.4.2), (C.4.4), (C.6.)*** *and* ***(C.7.)*** *are fulfilled. For all $0 < \widetilde{b}_0 < 1$, we infer that, as $n \to \infty$,*

$$\sup_{\varphi\widetilde{K} \in \mathscr{F}_m \mathscr{K}^m} \sup_{\mathbf{h} \in \mathcal{H}_0^{(m)}} \sup_{\mathbf{t} \in S_{\mathcal{X}}^m} \left|u_n(\varphi, \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right). \tag{3.15}$$

The following result handles the uniform deviation of the estimate $\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h})$ with respect to $\widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h})\right)$ in both situations, where the class of functions is bounded or unbounded satisfying a general moment condition.

Theorem 3.10. *Suppose that the conditions **(H.1.)**, **(C.3.1)**, **(C.4.1)**, **(C.4.4)**, **(C.6.)** and **(C.7.)** (or the following **(H.1.)**, **(C.3.1)**, **(C.4.2)**, **(C.4.4)**, **(C.6.)** and **(C.7.)**) are fulfilled. For all* $0 < \widetilde{b}_0 < 1$, *we infer, as* $n \to \infty$,

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{\mathbf{h}\in\mathcal{H}_0^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}) - \widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h})\right)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right). \tag{3.16}$$

Theorem 3.11. *Suppose that the conditions **(H.1.)**, **(C.2.1)**, **(C.3.1)** and **(C.6.)** are fulfilled. For all* $0 < \widetilde{h}'_n < 1, \widetilde{h}'_n \to 0$, *we infer, as* $n \to \infty$,

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h})\right) - r^{(m)}(\varphi,\mathbf{t})\right| = O\left(\widetilde{h'}_n^{\gamma}\right). \tag{3.17}$$

Corollary 3.12. *Under the assumptions of Theorems 3.10 and 3.11 it follows that, as* $n \to \infty$,

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}) - r^{(m)}(\varphi,\mathbf{t})\right| = O\left(\widetilde{h'}_n^{\gamma}\right) + O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right). \tag{3.18}$$

Remark 3.13. *As in [9, 27, 43], we will divide the U-statistics into different parts: some parts can be approximated by U-statistics of independent blocks, while by conditioning on one block, others are empirical processes of independent blocks. To prove the nonlinear terms are negligible, we will also need some symmetrization and maximal inequalities, we refer to [10, 67].*

3.4. Uniform consistency and UINN consistency of functional conditional U-statistics

Let $0 < \mu^* \le \nu^* < \infty$ be some constants and $\rho_n^* \in (0,1)$ is a sequence chosen in such a way that

$$\min_{1\le j\le m}\mu_j\phi^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right) = \mu^*\phi^{-1}\left(\frac{\rho_n^*k_{1,n}}{n}\right)$$

and

$$\max_{1\le j\le m}\nu_j\phi^{-1}\left(\frac{k_{2,n}}{\rho_{n,j}n}\right) = \nu^*\phi^{-1}\left(\frac{k_{2,n}}{\rho_n^*n}\right).$$

The following result deals with the uniform deviation of the estimate $u_n(\varphi,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t}))$ with respect to $\mathbb{E}\left(u_n(\varphi,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t}))\right)$ when the class of functions is bounded.

Theorem 3.14. *Suppose that the conditions **(H.1.)**, **(C.3.1)**, **(C.4.1)**, **(C.4.4)**,**(C.6.)** and **(C.7.)** are fulfilled. If in addition assumption **(C.8.)** holds, we infer that, as* $n \to \infty$,

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{k_{1,n}\le k\le k_{2,n}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|u_n(\varphi,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t})) - \mathbb{E}\left(u_n(\varphi,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t}))\right)\right|$$

$$= O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi^{-1}\left(\frac{\rho_n^*k_{1,n}}{n}\right)\right)}}\right). \tag{3.19}$$

The following result deals with the uniform deviation of the estimate $u_n(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t}))$ with respect to $\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t}))\right)$ when the class of functions is unbounded satisfying general moments condition.

Theorem 3.15. *Suppose that the conditions **(H.1.)**, **(C.3.1)**, **(C.4.2)**, **(C.4.4)**,**(C.6.)** and **(C.7.)** are fulfilled. If in addition assumption **(C.8.)** holds, we infer that, as $n \to \infty$,*

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{k_{1,n}\le k\le k_{2,n}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left| u_n(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t})) - \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t}))\right)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}\right). \tag{3.20}$$

The next results give uniform consistency when the class of functions is bounded or unbounded.

Theorem 3.16. *Suppose that the conditions **(H.1.)**, **(C.3.1)**, **(C.4.1)**, **(C.4.4)**, **(C.6.)** and **(C.7.)** (or the following **(H.1.)**, **(C.3.1)**, **(C.4.2)**, **(C.4.4)**, **(C.6.)** and **(C.7.)**) are fulfilled. If in addition assumption **(C.8.)** holds, we infer that, as $n \to \infty$,*

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{k_{1,n}\le k\le k_{2,n}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left| \widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t})) - \widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t}))\right)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}\right). \tag{3.21}$$

Theorem 3.17. *Suppose that the conditions **(H.1.)**, **(C.2.1)**, **(C.3.1)** and **(C.6.)** are fulfilled. If in addition assumption **(C.8.)** holds, we infer that, as $n \to \infty$,*

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{k_{1,n}\le k\le k_{2,n}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left| \widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t}))\right) - r^{(m)}(\varphi, \mathbf{t})\right| = O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)^{\gamma}\right). \tag{3.22}$$

Corollary 3.18. *Under the assumptions of Theorems 3.16 and 3.17 it follows that, as $n \to \infty$,*

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{k_{1,n}\le k\le k_{2,n}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left| \widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t})) - r^{(m)}(\varphi, \mathbf{t})\right| = O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)^{\gamma}\right) + O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}\right). \tag{3.23}$$

Remark 3.19. *The choice of the parameters μ^* and ρ_n^* defined in a similar way as in condition **(C.8.)** affects the rate of convergence of the k-NN estimator. We can choose these parameters depending on the small ball probability function $\phi_t(h)$.*

Remark 3.20. *The present work largely extends and completes the work of [42, 44] in several ways. There are basically no restrictions on the choice of the kernel function in our setup, apart from satisfying some mild conditions that we will give after. The selection of the bandwidth or the number of neighbors, however, is more problematic. It is worth noticing that the choice of bandwidth is crucial to obtain a good rate of consistency, for example, it has a big influence on the size of the estimate's bias. In general, we are interested in the selection of bandwidth and neighbors that produce an estimator that has a good balance between the bias and the variance of the considered estimators. It is then more appropriate to consider the bandwidth and neighbors varying according to the criteria applied and to the available data and location which cannot be achieved by using the classical methods. The interested reader may refer to [34, 146] for more details and discussion on the subject. In the present section, we have provided a response to this delicate problem in the FDA associated with the dependent data setting. In the present setting, we divide the U-statistics into different parts: some parts can be approximated by U-statistics of independent blocks, while by conditioning on one block, others are empirical processes of independent blocks. This decomposition is the key tool but makes the proof very involved which is the price of the extension to the dependent framework. This, allows us to use, in a nontrivial way, the techniques used for independent variables and mostly we will be using the results of [44]. We highlight that in the present paper, we have used a novel exponential inequality of [107] tailored to the dependent framework.*

4. Uniform central limit theorems

4.1. kNN conditional empirical process

We define the functional conditional empirical process for univariate bandwidth h_K by:

$$\left\{\nu_n(\psi; h_K \mid t) = \sqrt{k}\left(\widehat{r}_n^{*(1)}(\psi, t; h_K) - r^{(1)}(\psi, t)\right), \psi \in \mathscr{F}\mathscr{K}\right\}, \tag{4.1}$$

where $\widehat{r}_n^{*(1)}(\psi, t; h_K)$ designates (2.5) when $m = 1$, and $r^{(1)}(\psi, t)$ refers to the regression function (2.2), with

$$\psi(\cdot,\cdot) \in \mathscr{F}\mathscr{K} := \mathscr{F}_1\mathscr{K}^1 = \left\{\varphi(\cdot)K\left(\frac{d(\cdot, t)}{h_K}\right) : \varphi \in \mathscr{F}_1, K\left(\frac{d(\cdot, t)}{h_K}\right) \in \mathscr{K}^1\right\}.$$

If, for $\mathbb{P}\psi = \int \psi d\mathbb{P}$, where $\mathbb{P}$ is the probability measure and, for each $(\mathbf{x}, \mathbf{y})$,

$$\sup_{\psi\in\mathscr{F}\mathscr{K}} |\psi(\mathbf{x}, \mathbf{y}) - \mathbb{P}\psi| < \infty,$$

then $\{\nu_n(\psi \mid t) : \psi \in \mathscr{F}\mathscr{K}\}$ is a random element with values in $l_\infty(\mathscr{F}\mathscr{K})$, consisting of all functional ν_∞ on $\mathscr{F}\mathscr{K}$ such that

$$\sup_{\psi\in\mathscr{F}\mathscr{K}} |\nu_\infty(\psi)| < \infty.$$

Then, it will be important to investigate the following weak convergence

$$\{\nu_n(\psi \mid t) : \psi \in \mathscr{F}\mathscr{K}, t \in \mathbb{I}\} \xrightarrow{w} \{\mathbb{G}(\psi) : \psi \in \mathscr{F}\mathscr{K}\} \quad \text{in } l_\infty(\mathscr{F}\mathscr{K}).$$

It is known that the weak convergence to a Gaussian limit with a version of uniformly bounded and uniformly continuous paths (with respect to the $\|\cdot\|_2$) is equivalent to the finite-dimensional convergence and the existence of pseudo-metric $d_{\text{p.m}}$ on $\mathscr{F}\mathscr{K}$ such that $(\mathscr{F}\mathscr{K}, d_{\text{p.m}})$ is totally bounded pseudo-metric space and

$$\lim_{r\to 0}\limsup_{n\to\infty}\mathbb{P}^*\left\{\sup_{d_{\text{p.m}}(\psi_1,\psi_2)\leqslant r}|\nu_n((\psi_1-\psi_2)\mid t)|>\varepsilon\right\}=0. \tag{4.2}$$

Below, we write $Z \overset{d}{=} \mathcal{N}(\mu, \Sigma^2)$ whenever the random vector Z follows a normal law with vector expectation μ and matrix variance Σ^2, $\overset{d}{\to}$ denotes the convergence in distribution. The following theorem is adapted from [147] to the setting of the k-NN estimators. The main objective of this section is to investigate the central limit theorems for the functional conditional empirical process defined by

$$\left\{\nu_n(\psi; H_{n,k}(t)) \mid t) = \sqrt{k}\left(\widehat{r}_n^{*(1)}(\psi, t; H_{n,k}(t)) - r^{(1)}(\psi, t)\right), \psi \in \mathscr{F}\mathscr{K}\right\}. \tag{4.3}$$

Theorem 4.1. *Let consider the class of functions $\mathscr{F}\mathscr{K}$, suppose that conditions **(C.1'.)**, **(C.1.2)**, **(C.2.1)**, **(C.2.2)**, **(C.3'.)**, **(C.5.)** and **(C.8.)** hold. and if the smoothing parameter k satisfies for all $\gamma > 0$*

$$k\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^{2\gamma} \to 0 \text{ as } n \to \infty,$$

we get for: $\mathfrak{l} \geq 1 : \psi_1, \ldots, \psi_{\mathfrak{l}} \in \mathscr{F}\mathscr{K}$,

$$\{\nu_n(\psi_i; H_{n,k}(t)) \mid t) : i = 1, \ldots, \mathfrak{l}\} \overset{d}{\longrightarrow} \mathcal{N}(0, \Sigma),$$

where $\Sigma := (\sigma_{i,j})_{i,j=1,\ldots,\mathfrak{l}}$ is the covariance matrix with:

$$\sigma_{i,j} := \frac{\kappa_2' r^{(1)}(\varphi_i\varphi_j, t) - r^{(1)}(\varphi_i, t) r^{(1)}(\varphi_j, t)}{\kappa_1' f_1(t)}.$$

Theorem 4.2. *Suppose that the conditions **(C.3.1)**, **(C.4.2)–(C.5.1)**, and **(C.8.)** hold and for each $\varphi \in \mathscr{F}$,*

$$\mathbb{E}(\varphi^2(Y_1)) < \infty.$$

Then, we have

$$\lim_{b\to 0}\limsup_{n\to\infty}\mathbb{P}\left\{\sup_{\substack{\|\psi_1-\psi_2\|_p\leqslant b\\ \psi_1,\psi_2\in\mathscr{F}\mathscr{K}}}\left|\nu_n((\psi_1-\psi_2)\mid t)\right|>\varepsilon\right\}=0.$$

The two previous theorems can be summarized as follows:

Theorem 4.3. *Under conditions **(C.1'.)**, **(C.1.2)**, **(C.2.)**, **(C.3.1)**, **(C.3'.)**, **(C.4.4)** **(C.5.1)**, **(C.5.2)** and **(C.8.)**, then the process, as $n \to \infty$,*

$$\left\{\nu_n(\psi; H_{n,k}t)) \mid t) = \sqrt{k}\left(\widehat{r}_n^{*(1)}(\psi, t; H_{n,k}(t))) - r^{(1)}(\psi, t)\right) : \psi \in \mathscr{F}\mathscr{K}\right\},$$

converges in law to a Gaussian process $\{\mathbb{G}_n(\psi) : \psi \in \mathscr{F}\mathscr{K}\}$ that admits a version with uniformly bounded and uniformly continuous paths with respect to $\|\cdot\|_2$-norm.

Remark 4.4. *We mention that other types of applications can be obtained from Theorem 4.3 including the conditional distribution, conditional density, and conditional hazard function. This, and other applications of interest, will not be considered here due to lack of space.*

4.2. kNN conditional U-processes

In this section, we are interested in studying the weak convergence of conditional U-processes under absolute regular observations. Recall the class of functions considered is $\mathscr{F}_m\mathscr{K}^m$ given in Section 2. The conditional U-process indexed by $\mathscr{F}_m\mathscr{K}^m$:

$$\left\{U_n^{(m)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) := \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left(\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) - r^{(m)}(\varphi,\mathbf{t})\right)\right\}_{\mathscr{F}_m\mathscr{K}^m}, \tag{4.4}$$

The U-empirical process is defined by

$$\mu_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) := \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{u_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) - \mathbb{E}(u_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})))\right\}.$$

It should be noted that to establish the weak convergence of (4.4) it is first necessary to go through that of (4.6), below. Indeed, we will develop some details that will be used later. Because condition **(C.6.)** is satisfied, for each $\lambda > 0$, we have

$$\begin{aligned}
G_{\varphi,\mathbf{t},\mathbf{h}}(\mathbf{x},\mathbf{y}) &= G_{\varphi,\mathbf{t},\mathbf{h}}(\mathbf{x},\mathbf{y})\mathbb{1}_{\left\{\kappa^m F(\mathbf{y})\leq\lambda\left(n/\tilde{\phi}(\mathbf{h})\right)^{1/2(p-1)}\right\}} \\
&\quad + G_{\varphi,\mathbf{t},\mathbf{h}}(\mathbf{x},\mathbf{y})\mathbb{1}_{\left\{\kappa^m F(\mathbf{y})>\lambda\left(n/\tilde{\phi}(\mathbf{h})\right)^{1/2(p-1)}\right\}} \\
&=: G^{(T)}_{\varphi,\mathbf{t},\mathbf{h}}(\mathbf{x},\mathbf{y}) + G^{(R)}_{\varphi,\mathbf{t},\mathbf{h}}(\mathbf{x},\mathbf{y}).
\end{aligned} \tag{4.5}$$

We can write the U-statistic as follows

$$\begin{aligned}
\mu_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) &= \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{u_n^{(m)}\left(G^{(T)}_{\varphi,\mathbf{t},\mathbf{h}}\right) - \mathbb{E}\left(u_n^{(m)}\left(G^{(T)}_{\varphi,\mathbf{t},\mathbf{h}}\right)\right)\right\} \\
&\quad + \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{u_n^{(m)}\left(G^{(R)}_{\varphi,\mathbf{t},\mathbf{h}}\right) - \mathbb{E}\left(u_n^{(m)}\left(G^{(R)}_{\varphi,\mathbf{t},\mathbf{h}}\right)\right)\right\} \\
&=: \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{u_n^{(T)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t})) - \mathbb{E}\left(u_n^{(T)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t}))\right)\right\} \\
&\quad + \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{u_n^{(R)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t})) - \mathbb{E}\left(u_n^{(R)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t}))\right)\right\} \\
&=: \mu_n^{(T)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) + \mu_n^{(R)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})).
\end{aligned} \tag{4.6}$$

We call the first term of the right side of (4.6) $\mu_n^{(T)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))$ truncated part and the second $\mu_n^{(R)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))$ remainder part. First we are interested in $\mu_n^{(T)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))$. An application of Hoeffding's decomposition gives

$$\begin{aligned}
u_n^{(T)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t})) &= \sum_{p=0}^{m}\frac{m!}{(m-p)!}u_n^{(p)}\left(\pi_{p,m}G^{(T)}_{\varphi,\mathbf{t},\mathbf{h}}\right) \\
&= \mathbb{E}G^{(T)}_{\varphi,\mathbf{t},\mathbf{h}}(\mathbf{X}',\mathbf{Y}') + \sum_{p=1}^{m}\frac{m!}{(m-p)!}u_n^{(p)}\left(\pi_{p,m}G^{(T)}_{\varphi,\mathbf{t},\mathbf{h}}\right),
\end{aligned} \tag{4.7}$$

where $\{(\mathbf{X}'_i,\mathbf{Y}'_i)\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. r.v. with $\mathcal{L}\left(\mathbf{X}'_i,\mathbf{Y}'_i\right) = \mathcal{L}\left(\mathbf{X}_i,\mathbf{Y}_i\right)$ for each i, and $\mathbf{X}'$ and $\mathbf{Y}'$ are respectively defined as $\mathbf{X}$ and $\mathbf{Y}$. In view of (4.7), we have

$$\mu_n^{(T)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) = \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{\mathbb{E}G^{(T)}_{\varphi,t}(\mathbf{X}',\mathbf{Y}') + \sum_{p=1}^{m}\frac{m!}{(m-p)!}u_n^{(p)}\left(\pi_{p,m}G^{(T)}_{\varphi,\mathbf{t},\mathbf{h}}\right) - \mathbb{E}\left(u_n^{(T)}(\varphi,\mathbf{t})\right)\right\},$$

the stationarity assumption and some algebras show that

$$\mathbb{E}\left(u_n^{(T)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t}))\right)=\mathbb{E}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}(\mathbf{X}',\mathbf{Y}').$$

Therefore,

$$\begin{aligned}\mu_n^{(T)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t})) &= \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{\sum_{p=1}^{m}\frac{m!}{(m-p)!}u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}\right)\right\}\\ &= \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{mu_n^{(1)}\left(\pi_{1,m}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}\right)+\sum_{p=2}^{m}\frac{m!}{(m-p)!}u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}\right)\right\}. \qquad (4.8)\end{aligned}$$

By the fact that $\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}$ is $\mathbb{P}$-canonical, we have to show that

$$\left(\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\sum_{p=2}^{m}u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}\right)\right)\xrightarrow{\mathbb{P}}0.$$

So that to establish the weak convergence of the U-process $\left\{\mu_n^{(T)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))\right\}_{\mathscr{F}_m\mathscr{K}^m}$, it is enough to show

$$m\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}u_n^{(1)}\left(\pi_{1,m}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}\right)\xrightarrow{w}\mathbb{G}(\varphi)\quad\text{in }\ \ell_\infty\left(m\mathcal{G}^{(1)}\right),$$

where $\{\mathbb{G}(\varphi)\}_{m\mathcal{G}^{(1)}}$ is a Gaussian process indexed by $m\mathcal{G}^{(1)}$, and for $2\le p\le m$

$$\left\|\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}^{(T)}\right)\right\|_{\mathscr{F}_m\mathscr{K}^m}\xrightarrow{\mathbb{P}}0.$$

We have to prove after, that the remaining part is negligible, in the sense that

$$\left\|\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{u_n^{(R)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t}))-\mathbb{E}\left(u_n^{(R)}(\varphi,\mathbf{t},\mathbf{h}(\mathbf{t}))\right)\right\}\right\|_{\mathscr{F}_m\mathscr{K}^m}\xrightarrow{\mathbb{P}}0.$$

Nevertheless, when we have to deal with finite-dimensional convergence, the truncation does not matter. Which means that establishing the finite-dimensional convergence of $\mu_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))$ is equivalent to establishing that of $\mu_n^{(T)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))$.

Theorem 4.5. ***(a)*** *Under conditions* ***(C.1'.), (C.1.2), (C.2.), (C.3'.), (C.5.1), (C.5.2)*** *and if $r^{(m)}(\varphi,\mathbf{t})$ is continuous at* $\mathbf{t}$, *then, as $n\to\infty$,*

$$\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left(\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))-\mathbb{E}\left(u_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))\right)\right)\xrightarrow{d}\mathcal{N}\left(0,\rho^2\right), \qquad (4.9)$$

where

$$\begin{aligned}\rho^2 &:= m^2\left(\sigma_{\mathbf{t}}^2(\varphi,\varphi)-2r^{(m)}(\varphi,\mathbf{t})\sigma_{\mathbf{t}}^2(\varphi,1)+\left(r^{(m)}(\varphi,\mathbf{t})\right)^2\sigma_{\mathbf{t}}^2(1,1)\right), \qquad (4.10)\\ \sigma_{\mathbf{t}}^2\left(\varphi_i,\varphi_j\right) &:= \lim_{n\to\infty}\tilde{\phi}(\mathbf{h})\mathbb{E}\left(\pi_{1,m}G_{\varphi_i,\mathbf{t},\mathbf{h}},\pi_{1,m}G_{\varphi_j,\mathbf{t},\mathbf{h}}\right).\end{aligned}$$

(b) *If, in addition, the smoothing parameter k satisfies the condition* ***(C.8.)****, then we have, as* $n \to \infty$*,*

$$\sqrt{n^{-m+1}k^m}\left(\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}_{n,k}(\mathbf{t})) - \mathbb{E}\left(u_n(\varphi,\mathbf{t};\mathbf{h}_{n,k}(\mathbf{t}))\right)\right) \xrightarrow{d} \mathcal{N}\left(0,\rho^2\right), \tag{4.11}$$

where ρ^2 is defined as in (4.10), with

$$\sigma_{\mathbf{t}}^2\left(\varphi_i,\varphi_j\right) := \lim_{n\to\infty}\left(\frac{k}{n}\right)^m \mathbb{E}\left(\pi_{1,m}G_{\varphi_i,\mathbf{t},\mathbf{h}}, \pi_{1,m}G_{\varphi_j,\mathbf{t},\mathbf{h}}\right).$$

Corollary 4.6. *Under conditions* ***(C.1'.), (C.1.2), (C.2.), (C.3'.), (C.5.)*** *and if* $n\tilde{h'}_n^{2\gamma}\tilde{\phi}(\mathbf{h}(\mathbf{t})) \to 0$ *as* $n \to \infty$*, then we infer that:*

$$\sqrt{n^{-m+1}k^m}\left\{\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}_{n,k}(\mathbf{t})) - r^{(m)}(\varphi,\mathbf{t}\right\} \xrightarrow{d} \mathcal{N}\left(0,\rho^2\right). \tag{4.12}$$

Theorem 4.7. *Under conditions* ***(C.1'.), (C.1.2), (C.2.), (C.3'.), (C.5.1), (C.5.2), (C.8.)*** *and* $n\tilde{h'}_n^{2\gamma}\tilde{\phi}(\mathbf{h}(\mathbf{t})) \to 0$ *as* $n \to \infty$*. Let* $\mathscr{F}_m\mathscr{K}^m$ *be a measurable VC-subgraph class of functions from* $(\mathcal{X}^m,\mathcal{Y}^m) \longrightarrow \mathbb{R}$ *such as condition* ***(C.4.2)*** *is satisfied and, if the* β*-coefficients of the mixing stationary sequence* $\{(X_i,Y_i)\}_{i\in\mathbb{N}^*}$ *fulfill:*

$$\beta_s s^r \longrightarrow 0, \text{ as } s \to \infty, \tag{4.13}$$

for some $r > 1$, then $\left\{U_n^{(m)}(\varphi,\mathbf{t};\mathbf{h}_{n,k}(\mathbf{t}))\right\}_{\mathscr{F}_m\mathscr{K}^m}$ converges in law to a Gaussian process $\{\mathbb{G}(\varphi)\}_{\mathscr{F}_m\mathscr{K}^m}$which has a version with uniformly bounded and uniformly continuous paths with respect to $\|\cdot\|_2$−norm.

Remark 4.8. *It is worth noting what is the price to pay by the nice features of the k-NN-based estimators: remembering that* $H_{n,k}(X)$ *is a random variable (which depends on* $(X_1,\ldots,X_n)$*), one should expect that additional technical difficulties will appear along the proofs of asymptotic properties. To fix the idea on this point, note that the random elements involved in (2.1),* $m = 1$*, can not be decomposed as sums of independent variables (as it is the case for instance with kernel-based estimators), and hence its treatment will need more sophisticated probabilistic developments than standard limit theorems for sums of i.i.d. variables. Also, the Hoeffding decomposition can not be applied directly to (2.1), which is the main tool for the study U-statistics. The first step in proving Theorem 4.7 is the extension of [27, 43] to the multivariate bandwidth problem. In addition, we have considered new applications: the set indexed conditional U-statistic, Kendall rank correlation coefficient and time series prediction from a continuous set of past values. Another delicate problem lies in the fact that some maximal inequalities and symmetrisation techniques of [10, 67] are not applicable directly in our framework, making the proof quit lengthy, in particular, the equicontinuity of the empirical processes.*

Remark 4.9. *It is straightforward to modify the proofs of our results to show that it remains true when the entropy condition is substituted by the bracketing condition: For some* $C_0 > 0$ *and* $v_0 > 0$*,*

$$\mathcal{N}_{[\,]}(\epsilon,\mathscr{F}_m\mathscr{K}^m,L_2(\mathbb{P})) \leq C_0\epsilon^{-v_0}, 0 < \epsilon < 1.$$

Refer to p. 270 of [190] for the definition of $\mathcal{N}_{[\,]}(\epsilon,\mathscr{F}_m\mathscr{K}^m,L_2(\mathbb{P}))$.

5. Some potential applications

Although only four examples will be given here, they stand as archetypes for a variety of problems that can be investigated in a similar way.

5.1. Set indexed conditional U-statistics

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 $C_1 \times \cdots \times C_m := \widetilde{\mathbb{C}}$ in a is a class of sets $\mathscr{C}^m$,

$$\mathbb{G}^{(m)}(C_1 \times \cdots \times C_m \mid \mathbf{t}) = \mathbb{E}\left(\prod_{i=1}^{m} \mathbb{1}_{\{Y_i \in C_i\}} \mid (X_1, \ldots, X_m) = \mathbf{t}\right) \text{ for } \mathbf{t} \in \mathcal{X}^m.$$

We define metric entropy with the inclusion of the class of sets $\mathscr{C}$. For each $\varepsilon > 0$, the covering number is defined as :

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

the quantity $\log(\mathcal{N}(\varepsilon, \mathscr{C}, \mathbb{G}^{(1)}(\cdot \mid x)))$ is called metric entropy with inclusion of $\mathscr{C}$ with respect to $\mathbb{G}^{(1)}(\cdot \mid x)$. The quantity $\log \mathcal{N}(\varepsilon, \mathscr{C}, \mathbb{G}^{(1)}(\cdot \mid x))$ is called metric entropy with inclusion of $\mathscr{C}$ with respect to $\mathbb{G}(\cdot \mid x)$. Estimates for such covering numbers are known for many classes, (see, e.g., [81]). We will often assume below that either $\log \mathcal{N}(\varepsilon, \mathscr{C}, \mathbb{G}^{(1)}(\cdot \mid x))$ or $\mathcal{N}(\varepsilon, \mathscr{C}, \mathbb{G}^{(1)}(\cdot \mid x))$ behave like powers of ε^{-1}: we say that the condition (R_γ) holds if

$$\log \mathcal{N}(\varepsilon, \mathscr{C}, \mathbb{G}^{(1)}(\cdot \mid x)) \le H_\gamma(\varepsilon), \text{ for all } \varepsilon > 0, \tag{5.1}$$

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 [164], it is worth noticing that the condition (5.1), $\gamma = 0$, 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 $\mathbb{R}^d$ ($d \ge 2$) fulfill the condition (5.1), $\gamma = (d-1)/2$. This and other classes of sets satisfying (5.1) with $\gamma > 0$ can be found in [81]. As a particular case of (2.5), we estimate $\mathbb{G}^{(m)}(C_1 \times \cdots \times C_m \mid \mathbf{t})$

$$\widehat{\mathbb{G}}_n^{(m)}(\widetilde{\mathbb{C}}, \mathbf{t}) = \frac{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} \prod_{j=1}^{m} \mathbb{1}_{\{Y_{i_j} \in C_j\}} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}. \tag{5.2}$$

One can apply Corollary 3.18 to infer that

$$\sup_{\widetilde{\mathbb{C}} \times \widetilde{K} \in \mathscr{C}^m \mathscr{K}^m} \ \sup_{k_{1,n} \le k \le k_{2,n}} \ \sup_{\mathbf{t} \in S_{\mathcal{X}}^m} \left| \widehat{\mathbb{G}}_n^{(m)}(\widetilde{\mathbb{C}}, \mathbf{t}) - \mathbb{G}^{(m)}(\widetilde{\mathbb{C}} \mid \mathbf{t}) \right| \longrightarrow 0 \qquad \text{a.co.} \tag{5.3}$$

Remark 5.1. *Another point of view is to consider the following situation, for a compact* $\mathbf{J} \subset \mathbb{R}^{dm}$,

$$\mathbf{G}^{(m)}(y_1, \ldots, y_m \mid \mathbf{t}) = \mathbb{E}\left(\prod_{i=1}^{m} \mathbb{1}_{\{Y_i \le y_i\}} \mid (X_1, \ldots, X_m) = \mathbf{t}\right) \text{ for } \mathbf{t} \in \mathcal{X}^m, (y_1, \ldots, y_m) \in \mathbf{J}.$$

Let $L(\cdot)$ be a distribution in $\mathbb{R}^d$ and $\widetilde{H}_{n,k}(t_i)$ is the number of neighborhoods associated with Y_is. One can estimate $\mathbf{G}^{(m)}(y_1, \ldots, y_m \mid \mathbf{t}) = \mathbf{G}^{(m)}(\mathbf{y} \mid \mathbf{t})$ by

$$\widehat{\mathbf{G}}_n^{(m)}(\mathbf{y}, \mathbf{t}) := \frac{\sum_{(i_1,\ldots,i_m)\in I(m,n)} L\left(\frac{t_1 - Y_{i_1}}{\widetilde{H}_{n,\widetilde{k}}(t_1)}\right) \cdots L\left(\frac{t_m - Y_{i_m}}{\widetilde{H}_{n,\widetilde{k}}(t_m)}\right) K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}{\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}.$$

One can use Corollary 3.18 to infer that, as $n \to \infty$,

$$\sup_{\widetilde{K}\in\mathscr{K}^m} \sup_{k_{1,n}\le k,\widetilde{k}\le k_{2,n}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \sup_{\mathbf{y}\in\mathbf{J}} \left|\widehat{\mathbf{G}}_n^{(m)}(\mathbf{y}, \mathbf{t}) - \mathbf{G}^{(m)}(\mathbf{y} \mid \mathbf{t})\right| \longrightarrow 0 \qquad \text{a.co.} \tag{5.4}$$

5.2. Kendall rank correlation coefficient

To test the independence of one-dimensional random variables Y_1 and Y_2 [127] proposed a method based on the U-statistic K_n with the kernel function :

$$\varphi\left((s_1, t_1), (s_2, t_2)\right) = \mathbb{1}_{\{(s_2-s_1)(t_2-t_1)>0\}} - \mathbb{1}_{\{(s_2-s_1)(t_2-t_1)\leqslant 0\}}. \tag{5.5}$$

Its rejection on the region is of the form $\left\{\sqrt{n}K_n > \gamma\right\}$, for more general tests, refer [29, 32]. In this example, we consider a multivariate case. To test the conditional independence of $\boldsymbol{\xi}, \boldsymbol{\eta} : Y = (\boldsymbol{\xi}, \boldsymbol{\eta})$ given X, we propose a method based on the conditional U-statistic :

$$\widehat{r}_n^{(2)}(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t})) = \frac{\sum_{i\neq j}^{n} \varphi\left(Y_i, Y_j\right) K\left(\frac{d(t_1, X_i)}{H_{n,k}(t_1)}\right) K\left(\frac{d(t_2, X_j)}{H_{n,k}(t_2)}\right)}{\sum_{i\neq j}^{n} K\left(\frac{d(t_1, X_i)}{H_{n,k}(t_1)}\right) K\left(\frac{d(t_2, X_j)}{H_{n,k}(t_2)}\right)},$$

where $\mathbf{t} = (t_1, t_2) \in \mathbb{I} \subset \mathbb{R}^2$ and $\varphi(\cdot)$ is Kendall's kernel (5.5). Suppose that $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are d_1 and d_2-dimensional random vectors respectively and $d_1 + d_2 = d$. Furthermore, suppose that $\mathrm{Y}_1, \ldots, \mathrm{Y}_n$ are observations of $(\boldsymbol{\xi}, \boldsymbol{\eta})$, we are interested in testing :

$$\mathrm{H}_0 : \boldsymbol{\xi} \text{ and } \boldsymbol{\eta} \text{ are conditionally independent given } X. \text{ vs } \mathrm{H}_a : \mathrm{H}_0 \text{ is not true.} \tag{5.6}$$

Let $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2) \in \mathbb{R}^d$ such as $\|\mathbf{a}\| = 1$ and $\mathbf{a}_1 \in \mathbb{R}^{d_1}, \mathbf{a}_2 \in \mathbb{R}^{d_2}$, and $\mathrm{F}(\cdot), \mathrm{G}(\cdot)$ be the distribution functions of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ respectively. Suppose $F^{a_1}(\cdot)$ and $G^{a_2}(\cdot)$ to be continuous for any unit vector $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2)$ where

$\mathrm{F}^{\mathbf{a}_1}(t) = \mathbb{P}\left(\mathbf{a}_1^\top \boldsymbol{\xi} < t\right)$ and $\mathrm{G}^{\mathbf{a}_2}(t) = \mathbb{P}\left(\mathbf{a}_2^\top \boldsymbol{\eta} < t\right)$ and $\mathbf{a}_1^\mathrm{T}$ means the transpose of the vector $\mathbf{a}_i, 1 \leqslant i \leqslant 2$. For $n = 2$, let $Y^{(1)} = \left(\boldsymbol{\xi}^{(1)}, \boldsymbol{\eta}^{(1)}\right)$ and $Y^{(2)} = \left(\boldsymbol{\xi}^{(2)}, \boldsymbol{\eta}^{(2)}\right)$ such as $\boldsymbol{\xi}^{(i)} \in \mathbb{R}^{d_1}$ and $\boldsymbol{\eta}^{(i)} \in \mathbb{R}^{d_2}$ for $i = 1, 2$, and :

$$\varphi^a\left(\mathrm{Y}^{(1)}, \mathrm{Y}^{(2)}\right) = \varphi\left(\left(\mathbf{a}_1^\top \boldsymbol{\xi}^{(1)}, \mathbf{a}_2^\top \boldsymbol{\eta}^{(1)}\right), \left(\mathbf{a}_1^\top \boldsymbol{\xi}^{(2)}, \mathbf{a}_2^\top \boldsymbol{\eta}^{(2)}\right)\right).$$

An application of Corollary 3.18 gives, as $n \to \infty$,

$$\sup_{\varphi \widetilde{K} \in \mathscr{F}_2 \mathscr{K}^2} \sup_{k_{1,n} \le k \le k_{2,n}} \sup_{\mathbf{t} \in S_{\mathcal{X}}^2} \left| \widehat{r}_n^{*(2)}(\varphi^a, \mathbf{t}; \mathbf{h}_{n,k}(\mathbf{t})) - r^{(2)}(\varphi^a, \mathbf{t}) \right| \longrightarrow 0, \quad a.co. \tag{5.7}$$

5.3. Discrimination problems

Now, we apply the results to the problem of discrimination described in Section 3 of [185], refer to also to [184]. We will use a similar notation and setting. Let $\varphi(\cdot)$ be any function taking at most finitely many values, say $1, \ldots, M$. The sets

$$A_j = \{(\mathbf{y}_1, \ldots, \mathbf{y}_k) : \varphi(\mathbf{y}_1, \ldots, \mathbf{y}_k) = j\}, \quad 1 \le j \le M,$$

then yield a partition of the feature space. Predicting the value of $\varphi(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)$ is tantamount to predicting the set in the partition to which $(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)$ belongs. For any discrimination rule g, we have

$$\mathbb{P}(g(\mathbf{X}) = \varphi(\mathbf{Y})) \le \sum_{j=1}^{M} \int_{\mathbf{t}:g(\mathbf{t})=j\}} \max \mathfrak{M}^j(\mathbf{t}) d\mathbb{P}(\mathbf{t}),$$

where

$$\mathfrak{M}^j(\mathbf{t}) = \mathbb{P}(\varphi(\mathbf{Y}) = j \mid \mathbf{X} = \mathbf{t}), \quad \mathbf{t} \in \mathcal{X}^m.$$

The above inequality becomes equality if

$$g_0(\mathbf{t}) = \arg \max_{1 \le j \le M} \mathfrak{M}^j(\mathbf{t}).$$

$g_0(\cdot)$ is called the Bayes rule, and the pertaining probability of error

$$\mathbf{L}^* = 1 - \mathbb{P}(g_0(\mathbf{X}) = \varphi(\mathbf{Y})) = 1 - \mathbb{E}\left\{\max_{1 \le j \le M} \mathfrak{M}^j(\mathbf{t})\right\},$$

is called the Bayes risk. Each of the above unknown functions $\mathfrak{M}^j$'s can be consistently estimated by one of the methods discussed in the preceding sections. Let, for $1 \le j \le M$,

$$\mathfrak{M}_n^j(\mathbf{t}) = \frac{\displaystyle\sum_{(i_1,\ldots,i_m) \in I(m,n)} \mathbb{1}\{\varphi(Y_{i_1}, \ldots, Y_{i_k}) = j\} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}{\displaystyle\sum_{(i_1,\ldots,i_m) \in I(m,n)} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}. \tag{5.8}$$

Set

$$g_{0,n}(\mathbf{t}) = \arg \max_{1 \le j \le M} \mathfrak{M}_n^j(\mathbf{t}).$$

Let us introduce

$$\mathbf{L}_n^* = \mathbb{P}(g_{0,n}(\mathbf{X}) \neq \varphi(\mathbf{Y})).$$

The discrimination rule $g_{0,n}(\cdot)$ is asymptotically Bayes' risk consistent, as $n \to \infty$,

$$\mathbf{L}_n^* \to \mathbf{L}^*.$$

This follows from Corollary 3.18 and the obvious relation

$$\left|\mathbf{L}^* - \mathbf{L}_n^*\right| \leq 2\mathbb{E}\left[\max_{1\leq j\leq M}\left|\mathfrak{M}_n^j(\boldsymbol{X}) - \mathfrak{M}^j(\boldsymbol{X})\right|\right].$$

5.4. Time series prediction from a continuous set of past values

Let $\{Z_n(t),\ t \in \mathbb{R}\}_{n\geq 1}$ denote a sequence of processes with value in $\mathbb{R}$. Let s denote a fixed positive real number. In this model, we suppose that the process is observed from $t = 0$ until $t = t_{\max}$, and assume without loss of generality that $t_{\max} = nT + s < \tau$. The method ensures splitting the observed process into n fixed-length segments. Let us denote each piece of the process by

$$X_i = \{Z(t), (i-1)T \leq t < iT\}.$$

The response value is therefore $Y_i = Z(iT + s)$, and this can be formulated as a regression problem:

$$\varphi(Z_1(\tau + s), \ldots, Z_k(\tau + s)) = r^{(k)}(Z_1(t), \ldots, Z_k(t)), \text{ for } \tau - T \leq t < \tau. \tag{5.9}$$

provided that we assume that a function of this kind, r, does not depend on i (which is satisfied if the process is stationary, for example). Because of this, when we get to time τ, we can use the following predictor, which is directly derived from our estimator, to predict the value that will be at time $\tau + s$

$$\widehat{r}_n^{(k)}(\varphi, \mathbf{z}; m_n) = \frac{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} \varphi(Z_{i_1}(\tau + s), \ldots, Z_{i_k}(\tau + s)) K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}{\displaystyle\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)},$$

where $\mathbf{z} = (z_1, \ldots, z_k) = \{(Z_1(t), \ldots, Z_k(t)), \text{ for } \tau - T \leq t < \tau\}$. Corollary 3.12 provides mathematical support for this nonparametric functional predictor and extends previous results in numerous ways in [56, 91]. Notice that this modelization encompasses a wide variety of practical applications, as this procedure allows for the consideration of a large number of past process values without being affected by the curse of dimensionality. We believe that our findings will find applications beyond the scope of this work, in particular, because many popular measures of dependence, such as distance covariance and the Hilbert-Schmidt independence criterion can be estimated using U-statistics.

In the next section, we provide more details about how some of the methodologies of a number of neighbor choices in the literature can be combined with our results.

6. 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 [15, 44, 45, 50, 105, 109, 166]. This parameter has to be selected suitably, either in the standard finite-dimensional case or in the infinite-dimensional framework to ensure good practical performances. However, according to our knowledge, such studies do not presently exist for treating a such general functional conditional U-statistic (unless the real case we could find in the paper of [79] a paragraph devoted to the selection of the number k). Nevertheless an extension of the leave-one-out cross-validation procedure allows to define, for any fixed $\mathbf{i} = (i_1, \ldots, i_m) \in I(m,n)$:

$$\widehat{r}^{*(m)}_{n,\mathbf{i}}(\varphi, \mathbf{t}, \mathbf{h}_{n,k}(\mathbf{t})) = \frac{\sum_{\mathbf{j}\in I_n^m(\mathbf{i})} \varphi(Y_{j_1}, \ldots, Y_{j_m}) K\left(\frac{d(t_1, X_{j_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{j_m})}{H_{n,k}(t_m)}\right)}{\sum_{\mathbf{j}\in I_n^m(\mathbf{i})} K\left(\frac{d(t_1, X_{j_1})}{H_{n,k}(t_1)}\right) \cdots K\left(\frac{d(t_m, X_{j_m})}{H_{n,k}(j_m)}\right)}, \tag{6.1}$$

where

$$I_n^m(\mathbf{i}) := \{\mathbf{j} \in I(m,n) \text{ and } \mathbf{j} \neq \mathbf{i}\} = I(m,n)\backslash\{\mathbf{i}\}.$$

The Eq (6.1) represents the leave-out-$(\mathbf{X_i}, \mathbf{Y_i})$ estimator of the functional regression and also could be considered as a predictor of $\varphi(\mathbf{Y_i})$. In order to minimize the quadratic loss function, we introduce the following criterion, we have for some (known) non-negative weight function $\mathcal{W}(\cdot)$:

$$CV(\varphi, k) := \sum_{\mathbf{i}\in I(m,n)} \left(\varphi(\mathbf{Y_i}) - \widehat{r}^{*(m)}_{n,\mathbf{i}}(\varphi, \mathbf{X_i}, \mathbf{h}_{n,k}(\mathbf{X_i}))\right)^2 \widetilde{\mathcal{W}}(\mathbf{X_i}), \tag{6.2}$$

where

$$\widetilde{\mathcal{W}}(\mathbf{t}) := \prod_{i=1}^m \mathcal{W}(t_i).$$

Following the ideas developed by [166], a natural way for choosing the bandwidth is to minimize the precedent criterion, so let's choose $\widehat{k} \in [k_{1,n}, k_{2,n}]$ minimizing among $k \in [k_{1,n}, k_{2,n}]$:

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} CV(\varphi, k),$$

we can conclude, by Corollary 3.18, that, as $n \to \infty$,

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left|\widehat{r}^{*(m)}_{n,\mathbf{i}}(\varphi, \mathbf{t}, \mathbf{h}_{n,\widehat{k}}(\mathbf{t})) - r^{(m)}(\varphi, \mathbf{t})\right| \longrightarrow 0, \qquad \text{a.co.}$$

The main interest of our results is the possibility of deriving asymptotics for any automatic data-driven parameters. Let $K'(\cdot)$ be a density function in $\mathbb{R}^d$ and $H'_{n,k'}(t_i)$ is the number of neighborhoods associated with Y_is. One can estimate the conditional density $\mathbf{f}^{(m)}(y_1, \ldots, y_m \mid \mathbf{t}) = \mathbf{f}^{(m)}(\mathbf{y} \mid \mathbf{t})$ by

$$\widehat{\mathbf{f}}_n^{(m)}(\mathbf{y}, \mathbf{t})$$

$$:= \frac{\sum_{(i_1,\ldots,i_m)\in I(m,n)} K'\left(\frac{t_1 - Y_{i_1}}{H'_{n,k'}(t_1)}\right)\cdots K'\left(\frac{t_m - Y_{i_m}}{H_{n,k'}(t_m)}\right) K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}{\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{d(t_1, X_{i_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{i_m})}{H_{n,k}(t_m)}\right)}.$$

Hence, the leave-one-out estimator is given by

$$\widehat{\mathbf{f}}^{(m)}_{n,\mathbf{i}}(\mathbf{y},\mathbf{t}) := \frac{\sum_{(j_1,\ldots,j_m)\in I_n^m(\mathbf{i})} K'\left(\frac{t_1 - Y_{j_1}}{H'_{n,k'}(t_1)}\right)\cdots K'\left(\frac{t_m - Y_{j_m}}{H_{n,k'}(t_m)}\right) K\left(\frac{d(t_1, X_{j_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{j_m})}{H_{n,k}(t_m)}\right)}{\sum_{(i_1,\ldots,i_m)\in I(m,n)} K\left(\frac{d(t_1, X_{j_1})}{H_{n,k}(t_1)}\right)\cdots K\left(\frac{d(t_m, X_{j_m})}{H_{n,k}(t_m)}\right)}.$$

While the cross-validation procedures described above aim to approximate quadratic errors of estimation, alternative ways for choosing smoothing parameters could be introduced aiming rather to optimize the predictive power of the method. The criterion is given by

$$(\breve{k}, \breve{k}') = \arg \min_{k_{1,n}\le k\le k_{2,n},\ k'_{1,n}\le k'\le k'_{2,n}} \sum_{\mathbf{i}\in I(m,n)} \left(\varphi(\mathbf{Y_i}) - \arg\max_{\mathbf{y}\in\mathbf{J}} \widehat{\mathbf{f}}^{(m)}_{n,\mathbf{i}}(\mathbf{y},\mathbf{t})\right)^2.$$

7. Concluding remarks

In this paper, we consider the kNN kernel type estimator for conditional U-statistics, with the Nadaraya-Watson estimator as a special case, in a functional setting with regular datasets. To obtain our results, we need some regularity on the conditional U-statistics and conditional moments, decay rates on the probability of variables belonging to shrinking open balls, and convenient decreasing rates on mixing coefficients. In particular, the conditional moment assumption allows unbounded classes of functions to be considered. The proof of weak convergence adheres to a standard method: convergence of finite dimensions and the intricate equicontinuity of conditional U-processes. Approaching independence with a block decomposition technique, and then proving a central limit theorem for independent variables leads to finite-dimensional convergence. The equicontinuity requires more intricate control, and the details are lengthy due to the general and complex framework we have considered and will be presented in the following section. Observe that mixing is a type of asymptotic independence assumption that is commonly used to seek simplicity, but can be implausible when there is a strong dependence between the data. In [69] it is argued that β-mixing is the weakest mixing assumption that allows for a "complete" empirical process theory that incorporates maximal inequalities and uniform central limit theorems. There exist explicit upper bounds for β-mixing coefficients for Markov chains (cf. [113]) and for so-called V-geometric mixing coefficients (cf. [169]). For several stationary time series models like linear processes (cf. [202] for α-mixing), ARMA (cf. [189]), nonlinear AR (cf. [126]). A common assumption in these results is that the observed process or, more often, the innovations of the corresponding process, have a continuous distribution. This is a crucial assumption to handle the relatively complicated mixing coefficients defined over a supremum over two different sigma-algebras. A relaxation of β-mixing coefficients was investigated by ([103], Theorem 1) and is specifically designed for the analysis of the EDF, for more details, refer to [162]. The application of non-parametric functional concepts to general dependence structure is a relatively underdeveloped field. Notably, the ergodic framework eschews the commonly employed strong mixing condition and its variants for measuring dependence, as well as the extremely involved probabilistic calculations that this condition necessitates. It would be interesting to extend our work to the case of functional ergodic data, but this would require nontrivial mathematics and is well outside the scope of this paper. The primary obstacle lies in the necessity of formulating novel probabilistic results, as the ones employed in our current work, as demonstrated in [8], are tailored specifically for β-mixing samples. Another direction is to consider reducing the predictor's dimensionality by employing a Single Functional Index Model (SFIM) to estimate the regression [53]. SFIM has demonstrated its effectiveness in enhancing the consistency of the regression operator estimator. Change-point detection is widely employed to pinpoint positions within a data sequence where a stochastic system undergoes abrupt external influences. This method finds application across various scientific disciplines. The identification of these changes is crucial for exploring their diverse causes and enables appropriate responses. The challenge of detecting disruptions in a sequence of random variables has a rich historical background, refer to [26, 39, 41]. It would be of interest to find applications of our results in this direction.

8. Mathematical developments

This section is devoted to the proof of our results. The aforementioned notation is also used in what follows. The proof of Theorems are quite involved and will decomposed in several lemmas proved in Section A.

8.1. Proofs of uniform consistency results

8.1.1. General Lemma

We present Lemma 8.1 in a general setting, for instance, see [44], which could be useful in many other situations than ours; this is a generalization of a result obtained in [58]. More generally, this technical tool could be useful for dealing with random bandwidths.

Let $(\mathbf{A}_i, \mathbf{B}_i)_{n\geq 1}$ be n random vectors valued in $(\Omega^m \times \Omega'^m, \mathscr{A} \times \mathscr{B})$, a general space. Let S_Ω be a fixed subset of Ω and we note that $G : \mathbb{R} \times (S_\Omega \times \Omega) \to \mathbb{R}^+$ a function such that, $\forall t \in S_\Omega$, $G(\cdot, (t, \cdot))$ is measurable and $\forall x, x' \in \mathbb{R}$:

(L_0)

$$x \leq x' \Rightarrow G(x, \mathbf{z}) \leq G(x', \mathbf{z}) \ , \forall \mathbf{z} \in S_\Omega \times \Omega.$$

We define the pointwise measurable class of functions, for $1 \leq m \leq n$:

$$\mathscr{G}^m := \left\{ (x_1, \ldots, x_m) \mapsto \prod_{i=1}^{m} G(h_i, (x_i, t_i)), (h_1, \ldots, h_m) \in \mathbb{R}_+^m \backslash \{0\} \text{ and } (t_1, \ldots, t_m) \in S_\Omega^m \right\}.$$

Let $(\mathbf{D}_{n,k}(\mathbf{t}))_{n\in\mathbb{N}}$ be a sequence of random real vectors (r.r.v.) in such a way that for all $\mathbf{t} = (t_1, \ldots, t_m) \in S_\Omega^m$, $\mathbf{D}_{n,k}(\mathbf{t}) = (D_{n,k}(t_1), \ldots, D_{n,k}(t_m))$, and $\varphi : \Omega'^m \to \mathbb{R}$ be a measurable function belonging to some class of functions $\mathcal{F}_m$ and let $\mathfrak{M}^{(m)} : \mathcal{F}_m \times S_\Omega^m \to \mathbb{R}$ be a nonrandom function such that,

$$\sup_{\varphi\in\mathcal{F}_m}\sup_{\mathbf{t}\in S_\Omega^m}\left|\mathfrak{M}^{(m)}(\varphi,\mathbf{t})\right| < \infty.$$

Now, for all $\mathbf{t} \in S_\Omega^m$, $\varphi \in \mathcal{F}_m$ and $n \in \mathbb{N}\setminus\{0\}$ we define

$$\mathfrak{M}_n^{(m)}(\varphi,\mathbf{t};\mathbf{h}) = \frac{\sum_{i\in I(m,n)} \varphi(B_{i_1},\ldots,B_{i_m})\prod_{j=1}^m G\left(H_j,(t_j,A_{i,j})\right)}{\sum_{i\in I(m,n)}\prod_{j=1}^m G\left(H_j,(t_j,A_{i,j})\right)},$$

where $\mathbf{h} = (H_1,\ldots,H_m) \in \mathbb{R}_+^m$, and $\widetilde{G} = \prod_{i=1}^m G(h_i,(x_i,t_i))$, $\mathbf{t} = (t_1,\ldots,t_m) \in S_\Omega^m$.

Lemma 8.1. *Let $(\mathcal{U}_n)_{n\in\mathbb{N}}$ be a decreasing positive sequence such that $\lim_{n\to\infty}\mathcal{U}_n = 0$. If, for all increasing sequence $\xi_n \in (0,1)$ with $\xi_n - 1 = O(\mathcal{U}_n)$, there exist two sequences of r.r.v. $\left(\boldsymbol{D}_{n,k}^-(\xi_n,\mathbf{t})\right)_{n\in\mathbb{N}}$ and $\left(\boldsymbol{D}_{n,k}^+(\xi_n,\mathbf{t})\right)_{n\in\mathbb{N}}$ such that*

(L_1)

$$\forall n\in\mathbb{N} \text{ and } \mathbf{t}\in S_\Omega^m, D_{n,k}^-(\xi_n,t_j) \le D_{n,k}^+(\xi_n,t_j)\ ,\forall j = 1,\ldots,m,$$

(L_2)

$$\prod_{j=1}^m \mathbb{1}_{\left\{D_{n,k}^-(\xi_n,t_j)\le D_{n,k}(t_j)\le D_{n,k}^+(\xi_n,t_j)\right\}} \to 1,\quad a.co,\quad \forall\mathbf{t}\in S_\Omega^m.$$

(L_3)

$$\sup_{\widetilde{G}\in\mathscr{G}^m}\sup_{\mathbf{t}\in S_\Omega^m}\left|\frac{\sum_{i\in I(m,n)}\prod_{j=1}^m G\left(D_{n,k}^-(\xi_n,t_j),(t_j,A_{i,j})\right)}{\sum_{i\in I(m,n)}\prod_{j=1}^m G\left(D_{n,k}^+(\xi_n,t_j),(t_j,A_{i,j})\right)} - \xi_n\right| = O_{a.co}(\mathcal{U}_n),$$

(L_4)

$$\sup_{\varphi\widetilde{G}\in\mathscr{F}_m\mathscr{G}^m}\sup_{\mathbf{t}\in S_\Omega^m}\left|\mathfrak{M}_n^{(m)}\left(\varphi,\mathbf{t};\mathbf{D}_{n,k}^-(\xi_n,\mathbf{t})\right) - \mathfrak{M}^{(m)}(\varphi,\mathbf{t})\right| = O_{a.co}(\mathcal{U}_n),$$

(L_5)

$$\sup_{\varphi\widetilde{G}\in\mathscr{F}_m\mathscr{G}^m}\sup_{\mathbf{t}\in S_\Omega^m}\left|\mathfrak{M}_n^{(m)}\left(\varphi,\mathbf{t};\mathbf{D}_{n,k}^+(\xi_n,\mathbf{t})\right) - \mathfrak{M}^{(m)}(\varphi,\mathbf{t})\right| = O_{a.co}(\mathcal{U}_n).$$

Then, as $n\to\infty$,

$$\sup_{\varphi\widetilde{G}\in\mathscr{F}_m\mathscr{G}^m}\sup_{\mathbf{t}\in S_\Omega^m}\left|\mathfrak{M}_n^{(m)}(\varphi,\mathbf{t};\mathbf{D}_{n,k}(\mathbf{t})) - \mathfrak{M}^{(m)}(\varphi,\mathbf{t})\right| = O_{a.co}(\mathcal{U}_n). \tag{8.1}$$

We refer to [44] for proof of this lemma.

Proof of Theorem 3.1

In order to establish the convergence rates, the following notation is necessary. For all $t \in S_{\mathcal{X}}$, set

$$\begin{aligned}
\Delta_i(t; h_K(t)) &:= K\left(d(X_i, t)/h_K(t)\right), \\
\widehat{r}^{(1)}_{n,2}(\varphi, t; h_K(t)) &:= \frac{1}{n\mathbb{E}\left(\Delta_1(t; h_K(t))\right)} \sum_{i=1}^{n} \varphi(Y_i)\Delta_i(t; h_K(t)), \\
\widehat{r}^{(1)}_{n,1}(1, t; h_K(t)) &:= \frac{1}{n\mathbb{E}\left(\Delta_1(t; h_K(t))\right)} \sum_{i=1}^{n} \Delta_i(t; h_K(t)).
\end{aligned}$$

This allows us to write

$$\widehat{r}^{(1)}_{n}(\varphi, t; h_K(t)) = \widehat{r}^{(1)}_{n,2}(\varphi, t; h_K(t))/\widehat{r}^{(1)}_{n,1}(1, t; h_K(t)).$$

Now, let us consider the following decomposition

$$\begin{aligned}
\widehat{r}^{(1)}_{n}(\varphi, t; h_K(t)) - r^{(1)}(\varphi, t) &= \frac{1}{\widehat{r}^{(1)}_{n,1}(\varphi, t; h_K(t))}\left\{\widehat{r}^{(1)}_{n,2}(\varphi, t; h_K(t)) - \mathbb{E}\left[\widehat{r}^{(1)}_{n,2}(\varphi, t; h_K(t))\right]\right\} \\
&+ \frac{1}{\widehat{r}^{(1)}_{n,1}(\varphi, t; h_K(t))}\left\{\mathbb{E}\left[\widehat{r}^{(1)}_{n,2}(\varphi, t; h_K(t))\right] - r^{(1)}(\varphi, t)\right\} \\
&+ \frac{r^{(1)}(\varphi, t)}{\widehat{r}^{(1)}_{n,1}(\varphi, t; h_K(t))}\left\{1 - \widehat{r}^{(1)}_{n,1}(\varphi, t; h_K(t))\right\}.
\end{aligned}$$

Therefore, the proof of (8.2) is based on the following lemmas.

Lemma 8.2. *Under assumptions* ***(C.1.1), (C.3.1), (C.4.1), (C.5.1), (C.5.2'), (C.6.)*** *and* ***(C.7.)****, we have, as* $n \to \infty$,

$$\sup_{\varphi K \in \mathscr{F}\mathscr{K}} \sup_{h_{n,1} \le h_K \le h_{n,2}} \sup_{t \in S_{\mathcal{X}}} \left|\widehat{r}^{(1)}_{n,2}(\varphi, t, h_K(t)) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi, t, h_K(t))\right)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right), \tag{8.2}$$

and

$$\sup_{\varphi K \in \mathscr{F}\mathscr{K}} \sup_{h_{n,1} \le h_K \le h_{n,2}} \sup_{t \in S_{\mathcal{X}}} \left|\widehat{r}^{(1)}_{n,1}(\varphi, t, h_K(t)) - 1\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right). \tag{8.3}$$

This lemma gives us the rates of consistency of the stochastic part when the class of functions $\mathscr{F}$ is bounded. The following lemma will give us the result when the class of functions is unbounded.

Lemma 8.3. *Under assumptions the* ***(C.1.1), (C.3.1), (C.4.2), (C.5.1), (C.5.2'), (C.6.)*** *and* ***(C.7.)****, we have, as* $n \to \infty$,

$$\sup_{\varphi K \in \mathscr{F}\mathscr{K}} \sup_{h_{n,1} \leq h_K \leq h_{n,2}} \sup_{t \in S_{\mathcal{X}}} \left| \widehat{r}_{n,2}^{(1)}(\varphi, t, h_K(t)) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, h_K(t))\right) \right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}} \right). \quad (8.4)$$

Finally, we only need the following result for the bias term. This lemma can be obtained similar way as in [147], where more details are given.

Lemma 8.4. *Under the condition* **(C.2.1)**, *we have, as* $n \to \infty$,

$$\sup_{\varphi K \in \mathscr{F}\mathscr{K}} \sup_{h_{n,1} \leq h_K \leq h_{n,2}} \sup_{t \in S_{\mathcal{X}}} \left| \mathbb{E}\left[\widehat{r}_{n,2}^{(1)}(\varphi, t; h_K(t))\right] - r^{(1)}(\varphi, t) \right| = O(h_{n,2}^{\gamma}). \quad (8.5)$$

Proof of Theorem 3.3

Similar to [44], To prove Theorem 3.3 we need to check the conditions of Lemma 8.1 in the case of $m = 1$. For that, we first identify the variables as follows: $S_\Omega = S_{\mathcal{X}}$, $A_i = X_i$, $\varphi(B_i) = \varphi(Y_i)$,

$$\begin{aligned} G(H, (t, A_i)) &= K(H^{-1}d(t, X_i)), \\ D_{n,k}(t) &= H_{n,k}(t), \\ \mathfrak{M}_n^{(1)}(\varphi, t; H_{n,k}(t)) &= \widehat{r}_n^{*(1)}(\varphi, t; H_{n,k}(t)), \\ \mathfrak{M}(\varphi, t) &= r^{(1)}(\varphi, t). \end{aligned}$$

Choosing $D_{n,k}^-(\xi_n, t)$ and $D_{n,k}^+(\xi_n, t)$ such that

$$\phi_t(D_{n,k}^-(\xi_n, t)) = \frac{\sqrt{\xi_n} k}{n}, \quad (8.6)$$

$$\phi_t(D_{n,k}^+(\xi_n, t)) = \frac{k}{n\sqrt{\xi_n}}. \quad (8.7)$$

We denote $h^-(t) = D_{n,k}^-(\xi_n, t)$, $h^+(t) = D_{n,k}^+(\xi_n, t)$ and

$$\mathcal{U}_n = \phi^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_n} n}\right)^{\gamma} + \sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}},$$

for all increasing sequence $\xi_n \in (0, 1)$ such that $\xi_n - 1 = O(\mathcal{U}_n)$. Note that for all $\xi_n \in (0, 1)$, $t \in S_{\mathcal{X}}$ and $k_{1,n} \leq k \leq k_{2.n}$ we have

$$\phi_t^{-1}\left(\frac{\sqrt{\xi_n} k_{1,n}}{n}\right) \leq h^-(t) \leq \phi_t^{-1}\left(\frac{\sqrt{\xi_n} k_{2,n}}{n}\right),$$

$$\phi_t^{-1}\left(\frac{k_{1,n}}{n\sqrt{\xi_n}}\right) \leq h^+(t) \leq \phi_t^{-1}\left(\frac{k_{2,n}}{n\sqrt{\xi_n}}\right),$$

using the condition (2.12) we can easily deduce that the bandwidths $h^-(t)$ and $h^+(t)$ both belong to the interval

$$[h_{n,1}, h_{n,2}] = \left[\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right), \nu\phi^{-1}\left(\frac{k_{2,n}}{\rho_n n}\right)\right].$$

Checking the conditions (L_4) and (L_5)

Let us start with checking (L_4). The fact that ξ_n is bounded by 1, and the local bandwidth $h^-(t)$ satisfies the conditions of Theorem 3.1 gives

$$\begin{aligned}
&\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\leq h^-(t)\leq h_{n,2}}\ \sup_{t\in S_{\mathcal{X}}}\left|\mathfrak{M}_n^{(1)}\left(\varphi,t;D_{n,k}^-(\xi_n,t)\right)-\mathfrak{M}^{(1)}(\varphi,t)\right| \\
&= \sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\leq h^-(t)\leq h_{n,2}}\ \sup_{t\in S_{\mathcal{X}}}\left|\frac{\sum_{i=1}^{n}\varphi(Y_i)K\left(\frac{d(t,X_i)}{h^-(t)}\right)}{\sum_{i=1}^{n}K\left(\frac{d(t,X_i)}{h^-(t)}\right)}-r^{(1)}(\varphi,t)\right| \\
&= \sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\leq h^-(t)\leq h_{n,2}}\ \sup_{t\in S_{\mathcal{X}}}\left|\widehat{r}_n^{(1)}(\varphi,t;h^-(t))-r^{(1)}(\varphi,t)\right| \\
&= O_{a.co}\left(h_{n,2}^{\gamma}+O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right)\right),
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
&\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{k_{1,n}\leq k\leq k_{2,n}}\ \sup_{t\in S_{\mathcal{X}}}\left|\widehat{r}_n^{*(1)}(\varphi,t;h^-(t))-r^{(1)}(\varphi,t)\right| \\
&= O_{a.co}\left(\phi^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_n}n}\right)^{\gamma}+\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right) \\
&= O_{a.co}(\mathcal{U}_n).
\end{aligned}$$

We use the same reasoning to check (L_5) and we readily obtain

$$\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{k_{1,n}\leq k\leq k_{2,n}}\ \sup_{t\in S_{\mathcal{X}}}\left|\mathfrak{M}_n(\varphi,t;D_{n,k}^+(\xi_n,t))-\mathfrak{M}(\varphi,t)\right| = O_{a.co}(\mathcal{U}_n).$$

Hence, (L_4) and (L_5) are checked.

Checking the condition (L_2)

To check (L_2) we show that for all $t \in S_{\mathcal{X}}$ and $\varepsilon_0 > 0$,

$$\sum_{n\geq 1} \mathbb{P}\left\{\left|\mathbb{1}_{\left\{D^-_{n,k}(\xi_n,t)\leq H_{n,k}(t)\leq D^+_{n,k}(\xi_n,t)\right\}} - 1\right| > \varepsilon_0\right\} < \infty. \tag{8.8}$$

Let $\varepsilon_0 > 0$ be fixed. Let $\{t_1, \ldots, t_{N_\varepsilon(S_{\mathcal{X}})}\}$ be an ε-net for $S_{\mathcal{X}}$, for all $t \in S_{\mathcal{X}}$, we have

$$\begin{aligned}
&\mathbb{P}\left\{\left|\mathbb{1}_{\left\{D^-_{n,k}(\xi_n,t)\leq H_{n,k}(t)\leq D^+_{n,k}(\xi_n,t)\right\}} - 1\right| > \varepsilon_0\right\}\\
\leq\;& \mathbb{P}\left(H_{n,k}(t) \leq \phi_t^{-1}\left(\frac{\sqrt{\xi_n}k_{1,n}}{n}\right)\right) + \mathbb{P}\left(H_{n,k}(t) \geq \phi_t^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_n}n}\right)\right)\\
\leq\;& \sum_{\ell=1}^{N_{\varepsilon_n}(S_{\mathcal{X}})}\sum_{k=k_{1,n}}^{k=k_{2,n}} \mathbb{P}\left(H_{n,k}(t_\ell) \leq \phi_t^{-1}\left(\frac{\sqrt{\xi_n}k_{1,n}}{n}\right)\right)\\
&+ \sum_{\ell=1}^{N_{\varepsilon_n}(S_{\mathcal{X}})}\sum_{k=k_{1,n}}^{k=k_{2,n}} \mathbb{P}\left(H_{n,k}(t_\ell) \geq \phi_t^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_n}n}\right)\right)\\
\leq\;& N_{\varepsilon_n}(S_{\mathcal{X}})\sum_{k=k_{1,n}}^{k=k_{2,n}} \mathbb{P}\left(H_{n,k}(t_\ell) \leq \phi_t^{-1}\left(\frac{\sqrt{\xi_n}k_{1,n}}{n}\right)\right)\\
&+ N_{\varepsilon_n}(S_{\mathcal{X}})\sum_{k=k_{1,n}}^{k=k_{2,n}} \mathbb{P}\left(H_{n,k}(t_\ell) \geq \phi_t^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_n}n}\right)\right).
\end{aligned}$$

Now, we use a lemma similar to [124] (see Lemma B.5). For completeness, we give their proof. Making use of Lemma B.5, we infer that

$$\begin{aligned}
\mathbb{P}\left\{H_{n,k}(t) \leq \phi_t^{-1}\left(\frac{\alpha k_{1,n}}{n}\right)\right\} &= \mathbb{P}\left\{\sum_{i=1}^n \mathbb{1}_{B\left(t,\phi_t^{-1}\left(\alpha\frac{k_{1,n}}{n}\right)\right)}(X_i) > k\right\}\\
&= \mathbb{P}\left\{\sum_{i=1}^n \mathbb{1}_{B\left(t,\phi_t^{-1}\left(\alpha\frac{k_{1,n}}{n}\right)\right)}(X_i) > \frac{k}{\alpha k_{1,n}}\alpha k_{1,n}\right\}\\
&\leq \exp\left\{-\left(k - \alpha k_{1,n}\right)/4\right\}.
\end{aligned} \tag{8.9}$$

This implies that

$$\begin{aligned}
N_{\varepsilon_n}(S_{\mathcal{X}})\sum_{k=k_{1,n}}^{k_{2,n}} \mathbb{P}\left\{H_{n,k}(t) \leq \phi_t^{-1}\left(\frac{\alpha k_{1,n}}{n}\right)\right\} &\leq N_{\varepsilon_n}(S_{\mathcal{X}})k_{2,n}\exp\left\{-(1-\alpha)k_{1,n}/4\right\}\\
&\leq N_{\varepsilon_n}(S_{\mathcal{X}})n^{1-\{(1-\alpha)/4\}\frac{k_{1,n}}{\ln n}}.
\end{aligned}$$

In a similar way, we obtain

$$\mathbb{P}\left\{H_{n,k}(t) \geq \phi_t^{-1}\left(\frac{k_{2,n}}{\alpha n}\right)\right\} \leq \exp\left\{-\frac{\left(k_{2,n} - \alpha k\right)^2}{2\alpha k_{2,n}}\right\}. \tag{8.10}$$

It follows that

$$N_{\varepsilon_n}(S_{\mathcal{X}}) \sum_{k=k_{1,n}}^{k_{2,n}} \mathbb{P}\left\{H_{n,k}(t) \geq \phi_t^{-1}\left(\frac{k_{2,n}}{\alpha n}\right)\right\} \leq N_{\varepsilon_n}(S_{\mathcal{X}}) k_{2,n} \exp\{-(1-\alpha)k_{1,n}/2\alpha\}$$
$$\leq N_{\varepsilon_n}(S_{\mathcal{X}}) n^{1-\{(1-\alpha)/2\alpha\}\frac{k_{2,n}}{\ln n}}.$$

Therefore, by that fact that $k_{i,n}/\ln n \to \infty$, $i = 1, 2$, we obtain

$$N_{\varepsilon_n}(S_{\mathcal{X}}) \sum_{k=k_{1,n}}^{k_{2,n}} \mathbb{P}\left\{H_{n,k}(t) \leq \phi_t^{-1}\left(\frac{\alpha k_{1,n}}{n}\right)\right\} \leq N_{\varepsilon_n}(S_{\mathcal{X}}) n^{1-\{(1-\alpha)/4\}\frac{k_{1,n}}{\ln n}} < \infty, \tag{8.11}$$

$$N_{\varepsilon_n}(S_{\mathcal{X}}) \sum_{k=k_{1,n}}^{k_{2,n}} \mathbb{P}\left\{H_{n,k}(t) \geq \phi_t^{-1}\left(\frac{k_{2,n}}{\alpha n}\right)\right\} \leq N_{\varepsilon_n}(S_{\mathcal{X}}) n^{1-\{(1-\alpha)/2\alpha\}\frac{k_{2,n}}{\ln n}} < \infty. \tag{8.12}$$

Checking the condition (L_3)

We consider the following quantities:

$$Q_{n1} := \frac{\Delta_1(t, D_n^-(\xi_n, t))}{\Delta_1(t, D_n^+(\xi_n, t))},$$
$$Q_{n2} := \frac{\widehat{r}_{n,1}^{(1)}(\varphi, t; D_n^-(\xi_n, t))}{\widehat{r}_{n,1}^{(1)}(\varphi, t; D_n^+(\xi_n, t))} - 1,$$
$$Q_{n3} := \frac{\Delta_1(t, D_n^+(\xi_n, t))}{\Delta_1(t, D_n^-(\xi_n, t))}\xi_n - 1.$$

The condition (L_3) can be written as

$$\left| \frac{\sum_{i=1}^{n} K\left(\frac{d(t, X_i)}{D_n^-(\xi_n, t)}\right)}{\sum_{i=1}^{n} K\left(\frac{d(t, X_i)}{D_n^+(\xi_n, t)}\right)} - \xi_n \right| \leq |Q_{n1}|\,|Q_{n2}| + |Q_{n1}|\,|Q_{n3}|\,.$$

Hence, by the fact that $\xi_n \longrightarrow 1$, our claimed result is

$$\sup_{K\in\mathscr{K}} \sup_{k_{1,n}\leq k\leq k_{2,n}} \sup_{t\in S_{\mathcal{X}}} \left| \frac{\sum_{i=1}^{n} K\left(\frac{d(t, X_i)}{D_n^-(\xi_n, t)}\right)}{\sum_{i=1}^{n} K\left(\frac{d(t, X_i)}{D_n^+(\xi_n, t)}\right)} - \xi_n \right| = O_{a.co}(\mathcal{U}_n). \tag{8.13}$$

The proof of (8.13) is based on the following results

$$\sup_{K\in\mathscr{K}} \sup_{k_{1,n}\leq k\leq k_{2,n}} \sup_{t\in S_{\mathcal{X}}} |Q_{n1}| \leq C, \tag{8.14}$$

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}|Q_{n2}| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right), \tag{8.15}$$

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}|Q_{n3}| = O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n n}\right)^{\gamma}\right). \tag{8.16}$$

Proof of (8.14)

Using the condition **(C.3.1)** one has

$$\mathbb{E}\left(K\left(\frac{d(X_1,t)}{h(t)}\right)\right) \le \kappa_2\phi_t(h(t)).$$

Now using the condition **(C.1.1)** we directly obtain

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}|Q_{n1}| \le C. \tag{8.17}$$

Proof of (8.15)

We have

$$\begin{aligned}\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}|Q_{n2}| &= \sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}\left|\frac{Q_n(K,t,h^-(t))}{Q_n(K,t,h^+(t))}-1\right| \\ &\le \frac{1}{\inf_{K\in\mathscr{K}}\inf_{k_{1,n}\le k\le k_{2,n}}\inf_{t\in S_{\mathcal{X}}}\left|Q_n(K,t,h^+(t))\right|}\left(\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}\left|Q_n(K,t,h^-(t))-1\right|\right. \\ &\quad \left.+\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}\left|Q_n(K,t,h^+(t))-1\right|\right). \end{aligned}\tag{8.18}$$

To prove this, we use Lemma 8.2, which gives

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}\left|Q_n(K,t,h^-(t))-1\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right), \tag{8.19}$$

and

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}\left|Q_n(K,t,h^+(t))-1\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right). \tag{8.20}$$

Moreover, combining (8.19), (8.20) with the fact that

$$\sum_{n\ge1}\mathbb{P}\left(\inf_{K\in\mathscr{K}}\inf_{t\in S_{\mathcal{X}}}\widehat{r}_{n,1}^{(1)}(\varphi,t;D_n^+(\xi_n,t)) < C\right) < \infty, \tag{8.21}$$

it follows that

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\leq k\leq k_{2,n}}\sup_{t\in S_{\mathcal{X}}}|Q_{n2}| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu\phi^{-1}\left(\frac{\rho_n k_{1,n}}{n}\right)\right)}}\right). \quad (8.22)$$

Proof of (8.16)

[134] use in their proof of part the proof of Lemma 1 in [89], on the other, we will use some computations similar to the steps of the proof of Lemma 5.7.0.3 in [42]. Let us consider the following quantity:

$$\begin{aligned}
\hat{B}(t,\varphi K) &= \frac{\mathbb{E}[\widehat{r}^{(1)}_{n,2}(\varphi,t;D_n^+(\xi_n,t))]}{\mathbb{E}[\widehat{r}^{(1)}_{n,1}(\varphi,t;D_n^-(\xi_n,t))]} - r^{(1)}(\varphi,t),\\
\left|\hat{B}(t,\varphi K)\right| &= \left|\frac{\phi_t(D_n^-(\xi_n,t))}{\phi_t(D_n^+(\xi_n,t))}\frac{\mathbb{E}\left[K\left(\frac{d(t,X_1)}{D_n^+(\xi_n,t)}\right)\mathbb{E}[\varphi(Y_1)|X_1]\right]}{\mathbb{E}\left[K\left(\frac{d(t,X_1)}{D_n^-(\xi_n,t)}\right)\right]} - r^{(1)}(\varphi,t)\right|\\
&\leq \frac{1}{\mathbb{E}\left[K\left(\frac{d(t,X_1)}{D_n^-(\xi_n,t)}\right)\right]}\left|\frac{\phi_t(D_n^-(\xi_n,t))}{\phi_t(D_n^+(\xi_n,t))}\mathbb{E}\left[K\left(\frac{d(t,X_1)}{D_n^+(\xi_n,t)}\right)\left(\mathbb{E}[\varphi(Y)_1|X_1] - r^{(1)}(\varphi,t)\right)\right]\right|,
\end{aligned}$$

using the fact that $\frac{\phi_t(D_n^-(\xi_n,t))}{\phi_t(D_n^+(\xi_n,t))} = \xi_n$ and supposing that the condition **(C.2.1)** holds which means

$$\left|r^{(1)}(X,t) - r^{(1)}(\varphi,t)\right| \leq C_3 d^\gamma(X,t),$$

and assuming that the conditions (**C.1.1**) and (**C.3.1**) to be satisfied, then for all $t\in S_{\mathcal{X}}$, and $D_n^-(\xi_n,t)$ and $D_n^+(\xi_n,t)$ in $[h_{n,1},h_{n,2}]$, one gets

$$\begin{aligned}
\left|\frac{\mathbb{E}[\widehat{r}^{(1)}_{n,2}(\varphi,t;D_n^+(\xi_n,t))]}{\mathbb{E}[\widehat{r}^{(1)}_{n,1}(\varphi,t;D_n^-(\xi_n,t))]} - r^{(1)}(\varphi,t)\right| &\leq \frac{C_3\kappa_2\xi_n}{\kappa_1\phi_t\left(D_n^-(\xi_n,t)\right)}\left[\mathbb{E}\mathbb{1}_{B(t,D_n^+(\xi_n,t))}(X)d^\gamma(X,t)\right]\\
&\leq \frac{C_3\kappa_2\xi_n}{\kappa_1}\frac{\phi_t\left(D_n^+(\xi_n,t)\right)}{\phi_t\left(D_n^-(\xi_n,t)\right)}(D_n^+(\xi_n,t))^\gamma \leq \frac{C_2C_3\kappa_2\xi_n}{C_1\kappa_1}(D_n^+(\xi_n,t))^\gamma\\
&\leq C(D_n^+(\xi_n,t))^\gamma.
\end{aligned}$$

Keeping in mind the condition **(C.3.1)** and the fact that $\xi_n \longrightarrow 1$, we obtain

$$\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\sup_{(h^-(t),h^+(t))\in[h_{n,1},h_{n,2}]^2}\sup_{t\in S_{\mathcal{X}}}\left|\frac{\mathbb{E}[\widehat{r}^{(1)}_{n,2}(\varphi,t;D_n^+(\xi_n,t))]}{\mathbb{E}[\widehat{r}^{(1)}_{n,1}(\varphi,t;D_n^-(\xi_n,t))]} - r^{(1)}(\varphi,t)\right| \leq C'h_{n,2}^\gamma. \quad (8.23)$$

Finally, rewriting (8.23) with $\varphi\equiv 1$ gives us

$$\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\sup_{(h^-(t),h^+(t))\in[h_{n,1},h_{n,2}]^2}\sup_{t\in S_{\mathcal{X}}}\left|\frac{\mathbb{E}\left[K\left(\frac{d(t,X_1)}{D_n^+(\xi_n,t)}\right)\right]}{\mathbb{E}\left[K\left(\frac{d(t,X_1)}{D_n^-(\xi_n,t)}\right)\right]}\xi_n - 1\right| \leq C'h_{n,2}^\gamma,$$

which is equivalent to

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}|Q_{n3}| = O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n n}\right)^{\gamma}\right). \tag{8.24}$$

Combining the results of (8.17), (8.22) and (8.24) and the fact that $\xi_n \longrightarrow 1$, implies that

$$\sup_{K\in\mathscr{K}}\sup_{k_{1,n}\le k\le k_{2,n}}\sup_{t\in S_{\mathcal{X}}}\left|\frac{\sum_{i=1}^{n} K\left(\frac{d(t,X_i)}{D_n^-(\xi_n,t)}\right)}{\sum_{i=1}^{n} K\left(\frac{d(t,X_i)}{D_n^+(\xi_n,t)}\right)} - \xi_n\right| = O_{a.co}(\mathcal{U}_n).$$

Hence, (L_3) is checked. Note that (L_0) is obviously satisfied by **(C.3.1)**, and that (L_1) is also trivially satisfied by construction of $D_n^-(\xi_n,t)$ and $D_n^+(\xi_n,t)$. So one can apply Lemma 8.1, and (8.1) with $m = 1$ is exactly the result of Theorem 3.3.

Preliminaries of the proofs

This part is mainly dedicated to the study of the functional conditional U-statistics. Just like in the case of $m = 1$, where $S_{\mathcal{X}}$ is covered by

$$\bigcup_{\ell=1}^{N_\varepsilon(S_{\mathcal{X}})} B(t_\ell,\varepsilon),$$

for some radius ε. Hence, for each $\mathbf{t} \in S_{\mathcal{X}}^m$, there exists $\boldsymbol{\ell(t)} = (\ell(t_1),\ldots,\ell(t_m))$ where $\forall 1 \le i \le m, 1 \le \ell(t_i) \le N_\varepsilon(S_{\mathcal{X}})$ such that

$$\mathbf{t} \in \prod_{i=1}^{m} B(t_{\ell(t_i)},\varepsilon) \text{ and } d(t_i,t_{\ell(t_i)}) = \operatorname{argmin}_{1\le 1\le \ell(t_i)\le N_\varepsilon(S_{\mathcal{X}})} d(t_i,t_\ell).$$

So for each $\mathbf{t} \in S_{\mathcal{X}}^m$, the closest center is $\mathbf{t}_{\boldsymbol{\ell}(\mathbf{t})}$ and the ball with the closest center

$$\prod_{i=1}^{m} B(t_{\ell(t_i)},\varepsilon) := B(\mathbf{t}_{\boldsymbol{\ell}(\mathbf{t}),\varepsilon}).$$

The proofs of the UIB consistency for the multivariate bandwidth will follow the same lines as the proofs of the UIB consistency for the univariate smoothing parameter in [42, 44]. Furthermore, as in the proof of Theorem 3.1, we divide the sequence $\{(\mathbf{X}_i,\mathbf{Y}_i)\}$ into υ_n alternate blocks, here the sizes $a_n,\ b_n$ are different satisfying

$$b_n \ll a_n, \quad (\upsilon_n - 1)(a_n + b_n) < n \leqslant \upsilon_n(a_n + b_n), \tag{8.25}$$

and set, for $1 \leqslant j \leqslant \upsilon_n - 1$:

$$\begin{aligned}
H_j^{(U)} &= \{i : (j-1)(a_n+b_n)+1 \leqslant i \leqslant (j-1)(a_n+b_n)+a_n\},\\
T_j^{(U)} &= \{i : (j-1)(a_n+b_n)+a_n+1 \leqslant i \leqslant (j-1)(a_n+b_n)+a_n+b_n\},\\
H_{\upsilon_n}^{(U)} &= \{i : (\upsilon_n-1)(a_n+b_n)+1 \leqslant i \leqslant n \wedge (\upsilon_n-1)(a_n+b_n)+a_n\},\\
T_{\upsilon_n}^{(U)} &= \{i : (\upsilon_n-1)(a_n+b_n)+a_n+1 \leqslant i \leqslant n\}.
\end{aligned}$$

Proof of Theorem 3.8

In this section, we consider a bandwidth $\mathbf{h} = (h_1, \ldots, h_m) \in \mathcal{H}_n^{(m)}$. To prove Theorem 3.8, we can write the U-statistic for each $\mathbf{t} \in S_{\mathcal{X}}^m$ as follows

$$\begin{aligned}
&|u_n(\varphi, \mathbf{t}; \mathbf{h}) - \mathbb{E}[u_n(\varphi, \mathbf{t}; \mathbf{h})]| \\
&\leq \left|u_n(\varphi, \mathbf{t}; \mathbf{h}) - u_n(\varphi, \mathbf{t}_{\ell(\mathbf{t})}; \mathbf{h})\right| && (8.26)\\
&\quad + \left|\mathbb{E}[u_n(\varphi, \mathbf{t}_{\ell(\mathbf{t})}; \mathbf{h})] - \mathbb{E}[u_n(\varphi, \mathbf{t}; \mathbf{h})]\right| \\
&\quad + \left|u_n(\varphi, \mathbf{t}_{\ell(\mathbf{t})}; \mathbf{h}) - \mathbb{E}[u_n(\varphi, \mathbf{t}_{\ell(\mathbf{t})}; \mathbf{h})]\right| \\
&\leq \left|u_{n,11}(\varphi, \mathbf{t}; \mathbf{h})\right| + \left|u_{n,12}(\varphi, \mathbf{t}; \mathbf{h})\right| + \left|u_{n,13}(\varphi, \mathbf{t}; \mathbf{h})\right|. && (8.27)
\end{aligned}$$

Let us begin with the term $\left|u_{n,11}(\varphi, \mathbf{t}; \mathbf{h})\right|$, we have

$$\left|u_{n,11}(\varphi, \mathbf{t}; \mathbf{h})\right| \leq \frac{\mathbb{C}(n-m)!}{n!\tilde{\phi}(\mathbf{h})} \sum_{i \in I(m,n)} \left|\varphi(Y_{i_1}, \ldots, Y_{i_m}) \left\{\prod_{j=1}^m K\left(\frac{d(X_{ij}, t_j)}{h_j}\right) - \prod_{j=1}^m K\left(\frac{d(X_{ij}, t_{\ell(t_j)})}{h_j}\right)\right\}\right|.$$

By applying the Telescoping binomial, we get

$$\begin{aligned}
&\prod_{j=1}^m K\left(\frac{d(X_{ij}, t_j)}{h_j}\right) - \prod_{j=1}^m K\left(\frac{d(X_{ij}, t_{\ell(t_j)})}{h_j}\right) && (8.28)\\
&= \sum_{j=1}^m \left[\left\{K\left(\frac{d(X_{ij}, t_j)}{h_j}\right) - K\left(\frac{d(X_{ij}, t_{\ell(t_j)})}{h_j}\right)\right\}\right] \\
&\quad \times \prod_{q=1}^{j-1} K\left(\frac{d(X_{iq}, t_q)}{h_q}\right) \prod_{p=j+1}^m K\left(\frac{d(X_{ip}, t_{\ell(t_p)})}{h_p}\right). && (8.29)
\end{aligned}$$

From condition **(C.4.1)**, we could claim that

$$\prod_{q=1}^{j-1} K\left(\frac{d(X_{iq}, t_q)}{h_q}\right) \leq \kappa_2^{j-1} \prod_{q=1}^{j-1} \mathbb{1}_{B(t_q, h_q)}(X_{iq}),$$

similarly, we have

$$\prod_{p=j+1}^m K\left(\frac{d(X_{ip}, t_{\ell(t_p)})}{h_p}\right) \leq \kappa_2^{m-j} \prod_{p=j+1}^m \mathbb{1}_{B(t_{\ell(t_p)}, h_p)}(X_{ip}).$$

So, (8.28) satisfies :

$$\begin{aligned}
&\prod_{j=1}^m K\left(\frac{d(X_{ij}, t_j)}{h_j}\right) - \prod_{j=1}^m K\left(\frac{d(X_{ij}, t_{\ell(t_j)})}{h_j}\right) \\
&\qquad \prod_{q=1}^{j-1} \mathbb{1}_{B(t_q, h_q)}(X_{iq}) \prod_{p=j+1}^m \mathbb{1}_{B(t_{\ell(t_p)}, h_p)}(X_{ip})\Bigg] \\
&\leq \kappa_2^{m-1} \sum_{j=1}^m \left[\left\{K\left(\frac{d(X_{ij}, t_j)}{h_j}\right) - K\left(\frac{d(X_{ij}, t_{\ell(t_j)})}{h_j}\right)\right\} \mathbb{1}_{B(t_j, h_j) \cup B(t_{\ell(t_j)}, h_j)}(X_{ij})\right.
\end{aligned}$$

$$\prod_{q=1}^{j-1} \mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq}) \prod_{p=j+1}^{m} \mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\Bigg]$$

$$=: \sum_{j=1}^{m} \mathbf{K}_{i_j,h_j}(t_j, t_{\ell(t_j)}),$$

where

$$\begin{aligned}\mathbf{K}^{(\ell)}_{i_j,h_j}(t_j, t_{\ell(t_j)}) &= \kappa_2^{m-1}\left\{K\left(\frac{d(X_{ij}, t_j)}{h_j}\right) - K\left(\frac{d(X_{ij}, t_{\ell(t_j)})}{h_j}\right)\right\}\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\\ &\quad\times\prod_{q=1}^{j-1} \mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq}) \prod_{p=j+1}^{m} \mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip}).\end{aligned}$$

Therefore, we infer that

$$\begin{aligned}\left|u_{n,11}(\varphi,\mathbf{t};\mathbf{h})\right| &\leq \frac{\mathbb{C}(n-m)!}{n!\tilde{\phi}(\mathbf{h})}\kappa_2^{m-1}\sum_{i\in I(m,n)}\left|\varphi(Y_{i_1},\ldots,Y_{i_m})\sum_{j=1}^{m}\mathbf{K}^{(\ell)}_{i_j,h_j}(t_j, t_{\ell(t_j)})\right|\\ &\leq \frac{\mathbb{C}(n-m)!}{n!\tilde{\phi}(\mathbf{h})}M\kappa_2^{m-1}\sum_{i\in I(m,n)}\sum_{j=1}^{m}\left|\left\{K\left(\frac{d(X_{ij}, t_j)}{h_j}\right) - K\left(\frac{d(X_{ij}, t_{\ell(t_j)})}{h_j}\right)\right\}\right.\\ &\quad\left.\times\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\prod_{q=1}^{j-1} \mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq}) \prod_{p=j+1}^{m} \mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\right| \qquad (8.30)\\ &\leq \frac{(n-m)!}{n!}\mathbb{C}mM\kappa_2^{m-1}\sum_{i\in I(m,n)}\frac{1}{m}\sum_{j=1}^{m}\left[\frac{\varepsilon_n}{\tilde{\phi}(\mathbf{h})h_{n,j}}\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\right.\\ &\quad\left.\times\prod_{q=1}^{j-1} \mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq}) \prod_{p=j+1}^{m} \mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\right]. \qquad (8.31)\end{aligned}$$

The transition from Eq (8.30) to (8.31) is done thanks to the fact that the kernel function $K(\cdot)$ is Lipschitz. Uniformly on $\mathbf{t}\in S^m_{\mathcal{X}}$ and $\mathbf{h}\in\mathcal{H}^{(m)}_n$, we get

$$\begin{aligned}&\sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}^{(m)}_n}\sup_{\mathbf{t}\in S^m_{\mathcal{X}}}\left|u_{n,11}(\varphi,\mathbf{t};\mathbf{h})\right|\\ &\leq \sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}^{(m)}_n}\sup_{\mathbf{t}\in S^m_{\mathcal{X}}}\frac{(n-m)!}{n!}\sum_{i\in I(m,n)}\frac{1}{m}\sum_{j=1}^{m}\left[\frac{\mathbb{C}mM\kappa_2^{m-1}\varepsilon_n}{\tilde{\phi}(\mathbf{H}_n)h_{n,j}}\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\right.\\ &\quad\left.\times\prod_{q=1}^{j-1} \mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq}) \prod_{p=j+1}^{m} \mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\right]\\ &\leq \sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}^{(m)}_n}\sup_{\mathbf{t}\in S^m_{\mathcal{X}}}\frac{(n-m)!}{n!}\sum_{i\in I(m,n)}\frac{1}{m}\sum_{j=1}^{m}\left[\frac{\mathbb{C}mM\kappa_2^{m-1}\varepsilon_n}{\tilde{\phi}(\widetilde{\mathbf{H}}_{\mathbf{n}})\widetilde{h}_n}\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\right.\\ &\quad\left.\times\prod_{q=1}^{j-1} \mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq}) \prod_{p=j+1}^{m} \mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\right]\end{aligned}$$

$$\leq \sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\frac{(n-m)!}{n!}\sum_{i\in I(m,n)}\frac{1}{m}\sum_{j=1}^{m}\left[\frac{\mathbb{C}mM\kappa_2^{m-1}\varepsilon_n}{C_1'\phi(\widetilde{h}_n)\widetilde{h}_n}\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\right.$$
$$\left.\times\prod_{q=1}^{j-1}\mathbb{1}_{B(t_q,h_q)}(X_{iq})\prod_{p=j+1}^{m}\mathbb{1}_{B(t_{\ell(t_p)},h_p)}(X_{ip})\right],$$

by (3.2), where $\mathbf{H}_n=(h_{n,1},\ldots,h_{n,m})$ and $\widetilde{\mathbf{H}}_n=(\widetilde{h}_n,\ldots,\widetilde{h}_n)$, with $\widetilde{\mathbf{H}}_n\leq\mathbf{H}_n$ component by component. The idea is to apply Lemma B.6 on the function

$$f_{\mathbf{t},\mathbf{h}}(\mathbf{X_i})=\frac{1}{m}\sum_{j=1}^{m}\left[\frac{\mathbb{C}mM\kappa_2^{m-1}\varepsilon_n}{C_1'\phi(\widetilde{h}_n)h_{n,j}}\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\prod_{q=1}^{j-1}\mathbb{1}_{B(t_q,h_q)}(X_{iq})\prod_{p=j+1}^{m}\mathbb{1}_{B(t_{\ell(t_p)},h_p)}(X_{ip})\right],$$

which satisfies for all $\mathbf{t}\in S_{\mathcal{X}}^m$:

$$0\leq\sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\left|f_{\mathbf{t},\mathbf{h}}(\mathbf{X})\right|\leq\frac{\mathbb{C}mM\kappa_2^{m-1}\varepsilon_n}{C_1'\phi(\widetilde{h}_n)\widetilde{h}_n}\leq Cte=C_7.$$

Notice that the existence of the constant C_7 on the last right side of the preceding inequality is deduced from the condition (2.11). Now, we can apply Lemma B.6 with $x=\sqrt{\frac{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{n\phi(\widetilde{h}_n)}}-\frac{\varepsilon_n}{\widetilde{h}_n}$, which gives us

$$\mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|u_{n,11}(\varphi,\mathbf{t};\mathbf{h})\right|\geq\sqrt{\frac{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{n\phi(\widetilde{h}_n)}}\right\}\tag{8.32}$$
$$\leq\mathbb{P}\left\{\left|u_n^{(m)}(f)-\theta\right|\geq\sqrt{\frac{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{n\phi(\widetilde{h}_n)}}-\frac{\varepsilon_n}{\widetilde{h}_n}+C_{\gamma,m}/\sqrt{n}\right\}$$
$$\leq 2\exp\left\{-\frac{C_{\gamma,m}'x^2n}{C_7^2+C_7x(\log n)(\log_2(n))}\right\}\leq n^{-c\epsilon_0},\tag{8.33}$$

such that $c\epsilon_0>1$. By developing the computation while respecting the imposed conditions, mainly **(C.6.)** and **(C.7.)**, we get

$$\sum_{n\geq1}\mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|u_{n,11}(\varphi,\mathbf{t};\mathbf{h})\right|\geq\sqrt{\frac{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{n\phi(\widetilde{h}_n)}}\right\}<\infty.\tag{8.34}$$

The study of the term $u_{n,12}$ is deduced from the previous one. In fact:

$$\left|u_{n,12}(\varphi,\mathbf{t};\mathbf{h})\right|$$
$$=\left|\mathbb{E}\left[u_n(\varphi,\mathbf{t};\mathbf{h})-u_n(\varphi,\mathbf{t};\mathbf{h})\right]\right|$$
$$\leq\frac{\mathbb{C}(n-m)!}{n!\tilde{\phi}(\mathbf{h})}\left|\sum_{i\in I(m,n)}\mathbb{E}\left(\varphi(Y_{i_1},\ldots,Y_{i_m})\left\{\prod_{j=1}^{m}K\left(\frac{d(X_{ij},t_j)}{h_j}\right)-\prod_{j=1}^{m}K\left(\frac{d(X_{ij},t_{\ell(t_j)})}{h_j}\right)\right\}\right)\right|\tag{8.35}$$
$$\leq\frac{\mathbb{C}(n-m)!}{n!\tilde{\phi}(\mathbf{h})}\sum_{i\in I(m,n)}\mathbb{E}\left|\varphi(Y_{i_1},\ldots,Y_{i_m})\left\{\prod_{j=1}^{m}K\left(\frac{d(X_{ij},t_j)}{h_j}\right)-\prod_{j=1}^{m}K\left(\frac{d(X_{ij},t_{\ell(t_j)})}{h_j}\right)\right\}\right|.\tag{8.36}$$

To pass from (8.35) to (8.36), we apply Jensen's inequality in connection with some properties of the absolute value function. Then following the same way already taken, we get

$$\left|u_{n,12}(\varphi,\mathbf{t};\mathbf{h})\right| \leq \mathbb{C}\left|u'_{n,12}(\varphi,\mathbf{t};\mathbf{h})\right|,$$

where

$$\left|u'_{n,12}(\varphi,\mathbf{t};\mathbf{h})\right| := \frac{(n-m)!}{n!\tilde{\phi}(\mathbf{h})}\sum_{\mathbf{i}\in I(m,n)}\mathbb{E}\left|\varphi(Y_{i_1},\ldots,Y_{i_m})\sum_{j=1}^{m}\mathbf{K}^{(\ell)}_{i_j,h_j}(t_j,t_{\ell(t_j)})\right|.$$

Notice that we have

$$\begin{aligned}
\left|u'_{n,12}(\varphi,\mathbf{t};\mathbf{h})\right| &\leq \frac{(n-m)!}{n!\tilde{\phi}(\mathbf{h})}M\kappa_2^{m-1}\sum_{\mathbf{i}\in I(m,n)}\sum_{j=1}^{m}\mathbb{E}\left|\left\{K\left(\frac{d(X_{ij},t_j)}{h_j}\right)-K\left(\frac{d(X_{ij},t_{\ell(t_j)})}{h_j}\right)\right\}\right.\\
&\quad\left.\times\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\prod_{q=1}^{j-1}\mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq})\prod_{p=j+1}^{m}\mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\right|\\
&\leq \frac{(n-m)!}{n!}mM\kappa_2^{m-1}\sum_{\mathbf{i}\in I(m,n)}\frac{1}{m}\sum_{j=1}^{m}\left[\frac{\varepsilon_n}{\tilde{\phi}(\mathbf{H}_n)h_{n,j}}\mathbb{E}\left(\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\right.\right.\\
&\quad\left.\left.\times\prod_{q=1}^{j-1}\mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq})\prod_{p=j+1}^{m}\mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\right)\right]\\
&\leq \frac{(n-m)!}{n!}mM\kappa_2^{m-1}\sum_{\mathbf{i}\in I(m,n)}\frac{1}{m}\sum_{j=1}^{m}\left[\frac{\varepsilon_n}{C'_1\tilde{\phi}(\widetilde{h}_n)h_{n,j}}\mathbb{E}\left(\mathbb{1}_{B(t_j,h_j)\cup B(t_{\ell(t_j)},h_j)}(X_{ij})\right.\right.\\
&\quad\left.\left.\times\prod_{q=1}^{j-1}\mathbb{1}_{B\left(t_q,h_q\right)}(X_{iq})\prod_{p=j+1}^{m}\mathbb{1}_{B\left(t_{\ell(t_p)},h_p\right)}(X_{ip})\right)\right].
\end{aligned}$$

That implies

$$\begin{aligned}
\sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|u_{n,12}(\varphi,\mathbf{t};\mathbf{h})\right| &\leq mM\kappa_2^{m-1}\frac{\log(n)}{n\widetilde{h}_n}\leq C'_7\frac{\log(n)}{n\phi(\widetilde{h}_n)}\\
&= O\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{n\phi(\widetilde{h}_n)}}\right).
\end{aligned}$$

This gives that

$$\sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\sup_{\mathbf{h}\in\mathcal{H}_n^{(m)}}\sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|u_{n,12}(\varphi,\mathbf{t};\mathbf{h})\right| = O\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{n\phi(\widetilde{h}_n)}}\right). \tag{8.37}$$

Continue, now with $u_{n,13}$,

$$\left|u_{n,13}(\varphi,\mathbf{t};\mathbf{h})\right| = \left|u_n(\varphi,\mathbf{t}_\ell;\mathbf{h})-\mathbb{E}[u_n(\varphi,\mathbf{t}_\ell;\mathbf{h})]\right|.$$

Supposing that the kernel function $G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}(\cdot)$ is symmetric, we need to decompose our U-statistic according to [115] decomposition, we have

$$u_{n,13}(\varphi,\mathbf{t};\mathbf{h}) := u_n(\varphi,\mathbf{t}_\ell;\mathbf{h})-\mathbb{E}[u_n(\varphi,\mathbf{t}_\ell;\mathbf{h})]$$

$$= \sum_{p=1}^{m} \frac{m!}{(m-p)!} u_n^{(p)}\left(\pi_{p,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})\right)$$
$$= m u_n^{(1)}\left(\pi_{1,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})\right) + \sum_{p=2}^{m} \frac{m!}{(m-p)!} u_n^{(p)}\left(\pi_{p,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})\right). \quad (8.38)$$

Define new classes of functions, for $\mathbf{h} = (h_1, \ldots, h_m) \in \mathcal{H}_n^{(m)}$, $\boldsymbol{\ell} \in \{1, \ldots, N_{\varepsilon_n}(S_{\mathcal{X}})\}^m$, and $1 \le p \le m$:

$$(\mathscr{F}_m\mathscr{K}^m)^{(p)} := \left\{\phi_{\mathbf{t}_\ell}(\mathbf{h})\pi_{p,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})(\cdot,\cdot) \ \text{ for } \ \mathbf{h} \in \mathcal{H}_0^{(m)} \ \text{ and } \ \varphi \in \mathscr{F}_m\right\}.$$

These classes are VC-type classes of functions with the same characteristics and the envelope function F_p satisfying

$$F_p \le 2^p \kappa_2^m \| F\|_\infty.$$

Let us start with the linear term of (8.38), which is

$$m u_n^{(1)}\left(\pi_{1,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})\right) = \frac{m}{n} \sum_{j=1}^{m} \pi_{1,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})(X_i, Y_i).$$

From Hoeffding's projection, we have

$$\begin{aligned} \pi_{1,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})(x,y) &= \left\{\mathbb{E}\left[G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((x, X_2, \ldots, X_m), (y, Y_2, \ldots, Y_m)\right)\right] - \mathbb{E}[G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left(\mathbf{X}, \mathbf{Y}\right)]\right\} \\ &= \left\{\mathbb{E}[G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left(\mathbf{X}, \mathbf{Y}\right) | (X_1, Y_1) = (x,y)] - \mathbb{E}[G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left(\mathbf{X}, \mathbf{Y}\right)]\right\}. \end{aligned}$$

One can see that

$$m u_n^{(1)}\left(\pi_{1,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})\right) =: \frac{1}{\sqrt{n}} \alpha_n(S_{1,\mathbf{t}_\ell,\mathbf{h}}),$$

is an empirical process based on a VC-type class of functions contained in $m\,(\mathscr{F}_m\mathscr{K}^m)^{(1)}$ with the same characteristics and the elements are defined by:

$$S_{1,\mathbf{t}_\ell,\mathbf{h}}(x,y) = m\phi_{\mathbf{t}_\ell(\mathbf{h})}\mathbb{E}[G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left(\mathbf{X}, \mathbf{Y}\right) | (X_1, Y_1) = (x,y)].$$

Hence, the proof of this part is similar to that of the Lemma 8.2 and then:

$$\sup_{\mathbf{t} \in S_{\mathcal{X}}^m} \sup_{\widetilde{h}_n \le h_\ell \le b_0} \sup_{\varphi \in \mathscr{F}_m} \left| u_n^{(1)}\left(\pi_{1,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})\right)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right).$$

Pass now to the nonlinear terms. The purpose is to prove that, for $2 \le p \le m$:

$$\sup_{\varphi K \in \mathscr{F}_m\mathscr{K}^m} \sup_{\mathbf{h} \in \mathcal{H}_n^{(m)}} \sup_{\mathbf{t} \in S_{\mathcal{X}}^m} \frac{\binom{m}{p}\sqrt{n\phi(\widetilde{h}_n)}\left|u_n^{(p)}\left(\pi_{p,m}(G_{\varphi,\mathbf{t}_\ell,\mathbf{h}})\right)\right|}{\sqrt{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}} = O_{a.co}(1), \quad (8.39)$$

to do that, we need to decompose the interval $\prod_{j=1}^{m}(h_{n,j}, h'_{n,j})$ into smaller intervals. First, let us consider the intervals $\left(h_{n,j}, b_j\right)$ for all $j = 1, \ldots, m$ and $b_j \in (0, 1)$, we note

$$\mathcal{H}_{i_j} = \left[h_{i_j}, h'_{i_j}\right],$$

where $h_{i_j} = 2^{i-1}h_{n,j}$ and $h'_{i_j} = 2^{i}h_{n,j}$ and we set $L_j(n) = \max\left\{i : h'_{i_j} \leq 2b_j\right\}$, and $\mathcal{I}_j = \left\{i_j : 1 \leq i_j \leq L_j(n)\right\}$. We can observe that

$$\prod_{j=1}^{m}\left(h_{n,j}, b_j\right) \subseteq \bigcup_{(i_1,\ldots,i_m)\in\mathcal{I}_1\times\ldots\times\mathcal{I}_m} \prod_{j=1}^{m}\mathcal{H}_{i_j},$$

and

$$L_j(n) \sim \frac{\log(b_j/h_{n,j})}{\log(2)} \leq L(n) =: \max L_j(n) \text{ for } 1 \leq j \leq m. \tag{8.40}$$

Now, we set the following new classes, for $\mathbf{h} = (h_1, \ldots, h_m) \in \mathcal{H}_n^{(m)}$ and $\boldsymbol{\ell} \in \{1, \ldots, N_{\varepsilon_n}(S_{\mathcal{X}})\}^m$, $1 \leq j \leq m$, $1 \leq i \leq L(n)$ and $2 \leq p \leq m$:

$$\begin{aligned}
(\mathscr{F}_m\mathscr{K}^m)_{i_j,\boldsymbol{\ell}} &:= \left\{\phi_{\mathbf{t}_{\boldsymbol{\ell}}}(\mathbf{h})G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}(\cdot,\cdot) \text{ where } \varphi \in \mathscr{F}_m \text{ and } \mathbf{h} \in \prod_{j=1}^{m}\mathcal{H}_{i_j}\right\},\\
(\mathscr{F}_m\mathscr{K}^m)^{(p)}_{i_j,\boldsymbol{\ell}} &:= \left\{\phi_{\mathbf{t}_{\boldsymbol{\ell}}}(\mathbf{h})\pi_{p,m}(G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}})(\cdot,\cdot) \text{ where } \varphi \in \mathscr{F}_m \text{ and } \mathbf{h} \in \prod_{j=1}^{m}\mathcal{H}_{i_j}\right\}.
\end{aligned}$$

Thus, to prove (8.39), we need to prove that for $1 \leq j \leq m$ and $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_m)$:

$$\max_{1\leq\ell_j\leq N_{\varepsilon_n}(S_{\mathcal{X}})}\ \max_{1\leq i\leq L(n)}\ \sup_{\mathbf{h}\in\prod_{j=1}^{m}\mathcal{H}_{i_j}}\ \sup_{\varphi K\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \frac{\binom{m}{p}\sqrt{n\phi(\widetilde{h}_n)}\left|u_n^{(p)}\left(\pi_{p,m}(G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\right)\right|}{\sqrt{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}} = O_{a.co}(1).$$

Notice that for each ϵ_0, $1 \leq j \leq m$ and $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_m)$, we have

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{\mathbf{h}\in\mathcal{H}_0^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\ \sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \frac{\binom{m}{p}\sqrt{n\phi(\widetilde{h}_n)}\left|u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}(\mathbf{t})},\mathbf{h}}\right)\right|}{\sqrt{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}} \geq \epsilon_0\right\}\\
&\leq \sum_{1\leq\ell_j\leq N_{\varepsilon_n}(S_{\mathcal{X}})}\ \sum_{1\leq i\leq L(n)} \mathbb{P}\left\{\sup_{\mathbf{h}\in\prod_{j=1}^{m}\mathcal{H}_{i_j}}\ \sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \frac{\binom{m}{p}\sqrt{n\phi(\widetilde{h}_n)}\left|u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\right)\right|}{\sqrt{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}} \geq \epsilon_0\right\}
\end{aligned}$$

$$\leq L(n)N_{\varepsilon_n}^m(S_\mathcal{X}) \max_{1\leq \ell_j\leq N_{\varepsilon_n}(S_\mathcal{X})} \max_{1\leq i\leq L(n)}$$

$$\times \mathbb{P}\left\{\left\|\sum_{p=2}^{m}\binom{m}{p}u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t}_{\ell(\mathbf{t})},\mathbf{h}}\right)\right\|_{(\mathscr{F}_m\mathscr{K}^m)_{i_j,\ell}} \geq \epsilon_0\sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right\}.$$

At this stage, we will focus on studying the above equation for $m = 2$ to simplify the proof (the same steps remain valid for $m > 2$). We have

$$u_n^{(2)}\left(\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\right) = \frac{1}{n(n-1)}\sum_{1\leqslant<i<j\leqslant n}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,Y_i),(X_j,Y_j)\right),$$

and

$$\begin{aligned}\sum_{i\neq j}^{n}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right) &= \sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right)\\ &\quad+\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right)\\ &\quad+2\sum_{p=1}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right)\\ &\quad+2\sum_{p=1}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\leqslant 1}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right)\\ &\quad+\sum_{p\neq q}^{\upsilon_n}\sum_{i\in T_p^{(U)}}\sum_{j\in T_q^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right)\\ &\quad+\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in T_p^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right)\\ &=: \mathbb{I}+\mathbb{II}+\mathbb{III}+\mathbb{IV}+\mathbb{V}+\mathbb{VI}. \end{aligned} \quad (8.41)$$

For $m = 3$ the formula becomes cumbersome and given by

$$\begin{aligned}&\sum_{i\neq j\neq k}^{n}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)\\ &= \sum_{p_1\neq p_2\neq p_3}^{\upsilon_n}\sum_{i\in H_{p_1}^{(U)}}\sum_{j\in H_{p_2}^{(U)}}\sum_{k\in H_{p_3}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)\\ &\quad+3\sum_{p_1\neq p_2}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_{p_1}^{(U)}}\sum_{k\in H_{p_2}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)\end{aligned}$$

$$+\sum_{p_1=1}^{\upsilon_n}\sum_{i\neq j\neq k\, i,j,k\in H_{p_1}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1\neq p_2}^{\upsilon_n}\sum_{i\in H_{p_1}^{(U)}}\sum_{j\in H_{p_2}^{(U)}}\sum_{p_3:|p_3-p_i|\geqslant 2,i=1,2}^{\upsilon_n}\sum_{k\in T_{p_3}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1\neq p_2}^{\upsilon_n}\sum_{i\in H_{p_1}^{(U)}}\sum_{j\in H_{p_2}^{(U)}}\sum_{p_3:|p_3-p_1|\geqslant 2,|p_3-p_2|\leq 1}^{\upsilon_n}\sum_{k\in T_{p_3}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1=1}^{\upsilon_n}\sum_{i\in H_{p_1}^{(U)}}\sum_{p_3:|p_1-p_i|\geqslant 2,i=2,3}^{\upsilon_n}\sum_{j\in T_{p_2}^{(U)}}\sum_{k\in T_{p_3}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1=1}^{\upsilon_n}\sum_{i\in H_{p_1}^{(U)}}\sum_{p_3:|p_3-p_1|\geqslant 2,|p_2-p_1|\leq 1}^{\upsilon_n}\sum_{j\in T_{p_2}^{(U)}}\sum_{k\in T_{p_3}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1=1}^{\upsilon_n}\sum_{i,j\in H_{p_1}^{(U)}}\sum_{p_2:|p_2-p_1|\geqslant 2}^{\upsilon_n}\sum_{k\in T_{p_2}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1=1}^{\upsilon_n}\sum_{i\in H_{p_1}^{(U)}}\sum_{p_2:|p_2-p_1|\geqslant 2}^{\upsilon_n}\sum_{j,k\in T_{p_2}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1=1}^{\upsilon_n}\sum_{i\in H_{p_1}^{(U)}}\sum_{p_2:|p_2-p_1|\leq 1}^{\upsilon_n}\sum_{j,k\in T_{p_2}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1=1}^{\upsilon_n}\sum_{i,j\in H_{p_1}^{(U)}}\sum_{p_2:|p_2-p_1|\leqslant 1}^{\upsilon_n}\sum_{k\in T_{p_2}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+\sum_{p_1\neq p_2\neq p_3}^{\upsilon_n}\sum_{i\in T_{p_1}^{(U)}}\sum_{j\in T_{p_2}^{(U)}}\sum_{j\in T_{p_3}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+3\sum_{p_1\neq p_2}^{\upsilon_n}\sum_{i\in T_{p_1}^{(U)}}\sum_{j,k\in T_{p_2}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right)$$
$$+\sum_{p_1=1}^{\upsilon_n}\sum_{i\neq j\neq k\, i,j,k\in T_{p_1}^{(U)}}\pi_{\ell,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j,X_k),(Y_i,Y_j,Y_k)\right).$$

Let us start by considering the term $\mathbb{I}$. Suppose that the sequence of independent blocks $\{\xi_i=(\varsigma_i,\zeta_i)\}_{i\in\mathbb{N}^*}$ is of size a_n. An application of (A.1), shows that

$$\mathbb{P}\left\{\left\|\frac{1}{n(n-1)}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((X_i,X_j),(Y_i,Y_j)\right)\right\|_{\left(\mathscr{F}_2\mathscr{K}^2\right)_{i_j,\ell}}\geq\epsilon_0\sqrt{\frac{\psi_{S_X}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right\}$$

$$\leq \quad \mathbb{P}\left\{\left\|\frac{1}{n(n-1)}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\pi_{2,m}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right\|_{\left(\mathscr{F}_2\mathscr{K}^2\right)_{i_j,\ell}} \geq \epsilon_0\sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right\}$$
$$+2\upsilon_n\beta_{b_n}.$$

We keep the choice of υ_n and b_n such that $\upsilon_n {b_n}^r \leq 1$, which implies that $2\upsilon_n\beta_{b_n} \to 0$ as $n \to \infty$, so the term to consider is the second summand. the key idea is to apply Lemma B.4. We see clearly that the class of functions $(\mathscr{F}_m\mathscr{K}^m)_{i_j,\ell}$ is uniformly bounded, i.e.,

$$\sup_{\varphi\widetilde{K}\in\left(\mathscr{F}_2\mathscr{K}^2\right)_{i_j,\ell}}\left|\varphi(\cdot)\widetilde{K}(\cdot)\right| \leq M\kappa_2^2.$$

Moreover, by applying Proposition 2.6 of [10] we have for each $(x_i, y_i) \in S_\mathcal{X} \times \mathcal{Y}$ and Rademacher variables ϵ_i :

$$\begin{aligned}
&\mathbb{E}\left\|n^{-1}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right\|_{(\mathscr{F}_m\mathscr{K}^m)_{i_j,\ell}}\\
&\leqslant \quad c_2\mathbb{E}\left\|n^{-1}\sum_{p\neq q}^{\upsilon_n}\epsilon_p\epsilon_q\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right\|_{(\mathscr{F}_m\mathscr{K}^m)_{i_j,\ell}}\\
&\leqslant \quad c_2\mathbb{E}\int_0^{D_{2,n}}\log N\left(\epsilon,\left(\mathscr{F}_2\mathscr{K}^2\right)_{i_j,\ell},d_{2,n}\right)d\epsilon, \qquad (8.42)
\end{aligned}$$

where

$$D_{2,n} := \left\|\mathbb{E}_\epsilon^{1/2}\left\{n^{-2}\sum_{p\neq q}^{\upsilon_n}\left(\epsilon_p\epsilon_q\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right)^2\right\}\right\|_{(\mathscr{F}_m\mathscr{K}^m)_{i_j,\ell}}$$

and

$$d_{2,n} := \mathbb{E}_\epsilon^{1/2}\left\{n^{-2}\sum_{p\neq q}^{\upsilon_n}\left(\epsilon_p\epsilon_q\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right)^2\right\}.$$

We see that

$$D_{2,n} \leq n^{-1}\frac{n!}{(n-2)!}M\kappa_2^2 \leq nM\kappa_2^2.$$

So, using the fact that $\mathscr{F}_2\mathscr{K}^2$ is a VC-type class of functions satisfying **(C.4.4)** which implies that the class $\left(\mathscr{F}_2\mathscr{K}^2\right)_{i_j,\ell}$ is also a VC-type class of functions with the same characteristics as $\mathscr{F}_2\mathscr{K}^2$, then,

$$\mathbb{E}\left\|n^{-1}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}G_{\varphi,\mathbf{t}_\ell,\mathbf{h}}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right\|_{\left(\mathscr{F}_2\mathscr{K}^2\right)_{i_j,\ell}}$$

$$
\begin{aligned}
&\leq c_2 \mathbb{E} \int_0^{D_{2,n}} \log N\left(\epsilon, \left(\mathscr{F}_2 \mathscr{K}^2\right)_{i_j,\boldsymbol{\ell}}, d_{2,n}\right) d\epsilon \\
&\leq c_2 n M C^2. \qquad (8.43)
\end{aligned}
$$

All the conditions of Lemma B.4 are fulfilled, so a direct application gives for each ϵ_0,

$$
\begin{aligned}
&\mathbb{P}\left\{\left\|\frac{1}{n(n-1)} \sum_{p \neq q}^{\upsilon_n} \sum_{i \in H_p^{(U)}} \sum_{j \in H_q^{(U)}} \pi_{2,m} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right)\right\|_{\left(\mathscr{F}_2 \mathscr{K}^2\right)_{i_j,\boldsymbol{\ell}}} \geq \epsilon_0 \sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right\} \\
&\qquad \leq 2\exp\left(-\frac{\epsilon_0 \sqrt{\frac{2\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}}{2^7 2^3 c_2 M^2 \kappa_2^4}\right) \\
&\qquad \leq n^{-\varepsilon_0 C_2'}, \qquad (8.44)
\end{aligned}
$$

where

$$
\varepsilon_n = \frac{\log n}{n},\ C_2' > 0.
$$

Next, let us study the same blocks $\mathbb{II}$, we have

$$
\begin{aligned}
&\mathbb{P}\left\{\left\|\frac{1}{n(n-1)} \sum_{p=1}^{\upsilon_n} \sum_{i \neq j\, i,j \in H_p^{(U)}} \pi_{2,m} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right)\right\|_{\left(\mathscr{F}_2 \mathscr{K}^2\right)_{i_j,\boldsymbol{\ell}}} \geq \epsilon_0 \sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right\} \\
&\leq \mathbb{P}\left\{\left\|\frac{1}{n(n-1)} \sum_{p=1}^{\upsilon_n} \sum_{i \neq j\, i,j \in H_p^{(U)}} \pi_{2,m} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right)\right\|_{\left(\mathscr{F}_2 \mathscr{K}^2\right)_{i_j,\boldsymbol{\ell}}} \geq \epsilon_0 \sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right\} \\
&\quad + 2\upsilon_n \beta_{b_n}.
\end{aligned}
$$

Following the same argument as the blocks $\mathbb{I}$, we obtain

$$
\begin{aligned}
&\mathbb{E}\left\|n^{-1} \sum_{p=1}^{\upsilon_n} \sum_{i \neq j\, i,j \in H_p^{(U)}} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right)\right\|_{\left(\mathscr{F}_2 \mathscr{K}^2\right)_{i_j,\boldsymbol{\ell}}} \\
&\leqslant c_2 \mathbb{E}\left\|n^{-1} \sum_{p=1}^{\upsilon_n} \epsilon_p \sum_{i \neq j\, i,j \in H_p^{(U)}} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right)\right\|_{\left(\mathscr{F}_m \mathscr{K}^m\right)_{i_j,\boldsymbol{\ell}}}
\end{aligned}
$$

$$
\begin{aligned}
&\leq c_2 \mathbb{E} \int_0^{D_{2,n}^{(2)}} \log N\left(\epsilon, \left(\mathscr{F}_2 \mathscr{K}^2\right)_{i_j,\boldsymbol{\ell}}, d_{2,n}^{(2)}\right) d\epsilon \\
&\leq c_2 n M C'^2, \qquad (8.45)
\end{aligned}
$$

where

$$
D_{2,n}^{(2)} := \left\| \mathbb{E}_{\epsilon}^{1/2} \left\{ n^{-2} \left(\sum_{p=1}^{\upsilon_n} \epsilon_p \sum_{i \neq j\, i,j \in H_p^{(U)}} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right) \right)^2 \right\} \right\|_{(\mathscr{F}_m \mathscr{K}^m)_{i_j,\boldsymbol{\ell}}}
$$

and

$$
d_{2,n}^{(2)} := \mathbb{E}_{\epsilon}^{1/2} \left\{ n^{-2} \left(\sum_{p=1}^{\upsilon_n} \epsilon_p \sum_{i \neq j\, i,j \in H_p^{(U)}} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right) \right)^2 \right\}.
$$

Again, using Lemma B.4, we readily obtain

$$
\begin{aligned}
&\mathbb{P} \left\{ \left\| \frac{1}{n(n-1)} \sum_{p=1}^{\upsilon_n} \sum_{i \neq j\, i,j \in H_p^{(U)}} \pi_{2,m} G_{\varphi,\mathbf{t}_{\boldsymbol{\ell}},\mathbf{h}}\left((X_i, X_j), (Y_i, Y_j)\right) \right\|_{\left(\mathscr{F}_2 \mathscr{K}^2\right)_{i_j,\boldsymbol{\ell}}} \geq \epsilon_0 \sqrt{\frac{\psi_{S_X}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}} \right\} \\
&\leq 2 \exp\left(- \frac{\epsilon_0 \sqrt{\frac{2\psi_{S_X}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}}{2^7 2^3 c_2 M^2 \kappa_2^4} \right) \\
&\leq n^{-\varepsilon_0 C_3'}. \qquad (8.46)
\end{aligned}
$$

The results for the remaining blocks can be obtained by following the same strategy above. Consequently, we have

$$
\begin{aligned}
&\sum_{n \geq 1} \mathbb{P} \left\{ \sup_{\mathbf{h} \in \mathcal{H}_0^{(m)}} \sup_{\mathbf{t} \in S_{\mathcal{X}}^m} \sup_{\varphi \widetilde{K} \in \mathscr{F}_m \mathscr{K}^m} \frac{\binom{m}{p} \sqrt{n\phi(\widetilde{h}_n)} \left| u_n^{(p)}\left(\pi_{p,m} G_{\varphi,\mathbf{t}_{\ell(\mathbf{t})},\mathbf{h}}\right) \right|}{\sqrt{\psi_{S_X}\left(\frac{\log n}{n}\right)}} \geq \epsilon_0 \right\} \\
&\leq \sum_{n \geq 1} L(n) n^{-\varepsilon_0 C_m'} \\
&\leq \infty,
\end{aligned}
$$

where

$$
\varepsilon_0 = \frac{\log n}{n}, \; C_m' > 0.
$$

This completes the proof of the theorem. ■

Proof of Theorem 3.10:

Notice that

$$\begin{aligned}
&\left|\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}) - \widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h})\right)\right| \\
&= \left|\frac{u_n(\varphi, \mathbf{t}, \mathbf{h})}{u_n(1, \mathbf{t}, \mathbf{h})} - \frac{\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)}{\mathbb{E}\left(u_n(1, \mathbf{t}, \mathbf{h})\right)}\right| \\
&\leq \frac{\left|u_n(\varphi, \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right|}{\left|u_n(1, \mathbf{t}, \mathbf{h})\right|} \\
&\qquad + \frac{\left|\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right| \cdot \left|u_n(1, \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(1, \mathbf{t}, \mathbf{h})\right)\right|}{\left|u_n(1, \mathbf{t}, \mathbf{h})\right| \cdot \left|\mathbb{E}\left(u_n(1, \mathbf{t}, \mathbf{h})\right)\right|} \\
&=: \mathbf{I} + \mathbf{II}.
\end{aligned}$$

Under the imposed hypothesis and the previously obtained results, and for some c_1', c_2', we get that

$$\begin{aligned}
\sup_{\mathbf{h}\in\mathcal{H}_0^{(m)}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left|u_n(1, \mathbf{t}, \mathbf{h})\right| &= c_1' \qquad \text{a.co} \\
\sup_{\mathbf{h}\in\mathcal{H}_0^{(m)}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left|\mathbb{E}\left(u_n(1, \mathbf{t}, \mathbf{h})\right)\right| &= c_2' \\
\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{\mathbf{h}\in\mathcal{H}_0^{(m)}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \left|\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right| &= O(1).
\end{aligned}$$

Therefore, we can now apply Theorem 3.8 to handle **II** and Theorems 3.8 and 3.9 to handle **I**, depending on whether the class $\mathscr{F}_m$ satisfies **(C.4.1)** or **(C.4.2)**, we get, for some $c'' > 0$ with probability 1:

$$\begin{aligned}
&\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{\mathbf{h}\in\mathcal{H}_0^{(m)}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \frac{\sqrt{n\phi(\widetilde{h}_n)}\left|\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h}) - \widehat{\mathbb{E}}\left(\widehat{r}_n^{*(m)}(\varphi, \mathbf{t}; \mathbf{h})\right)\right|}{\sqrt{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}} \\
&\qquad \leq \sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{\mathbf{h}\in\mathcal{H}_0^{(m)}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \frac{\sqrt{n\phi(\widetilde{h}_n)}\,(\mathbf{I})}{\sqrt{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}} \\
&\qquad + \sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m} \sup_{\mathbf{h}(\mathbf{t})\in\mathcal{H}_0^{(m)}} \sup_{\mathbf{t}\in S_{\mathcal{X}}^m} \frac{\sqrt{n\phi(\widetilde{h}_n)}\,(\mathbf{II})}{\sqrt{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}} \\
&\qquad \leq c''.
\end{aligned}$$

Hence, the proof is complete.

Proof of Theorem 3.11:

Under the conditions **(C.3.1)**, we have

$$\begin{aligned}
&\left|\frac{\mathbb{E}[u_n(\varphi, \mathbf{t}; \mathbf{h})]}{\mathbb{E}[u_n(1, \mathbf{t}; \mathbf{h})]} - r^{(m)}(\varphi, \mathbf{t})\right| \\
&\quad \leq \quad \left|\kappa_1^{-m} C_1^{-1} \mathbb{E}[u_n(\varphi, \mathbf{t}; \mathbf{h})] - r^{(m)}(\varphi, \mathbf{t})\right|
\end{aligned}$$

$$
\begin{aligned}
&\leq \left|\kappa_1^{-m} C_1^{-1} \mathbb{E}\left(\frac{(n-m)!}{n!} \sum_{i \in I(m,n)} \frac{1}{\phi_{\mathbf{t}}(\mathbf{h})} \varphi(Y_{i_1}, \ldots, Y_{i_m}) \prod_{j=1}^{m} K\left(\frac{d(X_{i_j}, t_j)}{h_j}\right)\right)\right. \\
&\quad \left. - r^{(m)}(\varphi, \mathbf{t})\right| \\
&\leq \left|\kappa_1^{-m} C_1^{-1} \frac{1}{\phi_{\mathbf{t}}(\mathbf{h})} \mathbb{E}\left(\varphi(Y_1, \ldots, Y_m) \prod_{i=1}^{m} K\left(\frac{d(X_i, t_i)}{h_i}\right)\right) - r^{(m)}(\varphi, \mathbf{t})\right| \\
&\leq \frac{1}{\phi_{\mathbf{t}}(\mathbf{h}) \kappa_1^m C_1} \mathbb{E}\left(\prod_{i=1}^{m} K\left(\frac{d(X_i, t_i)}{h_i}\right) \left|r^{(m)}(\varphi, \mathbf{X}) - r^{(m)}(\varphi, \mathbf{t})\right|\right).
\end{aligned}
$$

Taking in consideration the hypotheses **(H.1)**, **(C.2.1)**, **(C.3.1)** and **(C.6.)**, we get $\forall \mathbf{h} = (h_1, \ldots, h_m) \in \prod_{j=1}^{m} (h_{n,j}, b_j)$ and $\forall \mathbf{t} \in S_{\mathcal{X}}^m$:

$$
\begin{aligned}
\left|\kappa_1^{-m} C_1^{-1} \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right) - r^{(m)}(\varphi, \mathbf{t})\right| &\leq \frac{C_3}{\phi_{\mathbf{t}}(\mathbf{h}) \kappa_1^m C_1} \mathbb{E}\left(\prod_{i=1}^{m} K\left(\frac{d(X_i, t_i)}{h_i}\right) d_{\mathcal{X}^m}^{\gamma}(\mathbf{X}, \mathbf{t})\right) \\
&\leq \frac{C_3}{m} \left(d(X_1, t_1) + \cdots + d(X_m, t_m)\right)^{\gamma} \\
&\leq \frac{C_3}{m} (h_1 + \cdots + h_m)^{\gamma} \\
&= \frac{C_3}{m} (b_1 + \ldots + b_m)^{\gamma} \leq C_3' \widetilde{b}_0^{\gamma},
\end{aligned}
$$

where

$$
\widetilde{b}_0 := \max_{1 \leq j \leq m} b_j.
$$

This completes the proof of the theorem.

Proof of Corollary 3.18:

In this section, we will prove Corollary 3.18 using Lemma 8.1. Following the same reasoning as the case of the functional regression, we use the notation: $S_{\Omega} = S_{\mathcal{X}}$, $\mathbf{A}_i = \mathbf{X}_i$, $\varphi(\mathbf{B}_i) = \varphi(\mathbf{Y}_i)$,

$$
\begin{aligned}
\prod_{i=1}^{m} G(H_i, (t, A_i)) &= \prod_{i=1}^{m} K(H_i^{-1} d(t, A_i)), \\
D_{n,k}(t_j) &= H_{n,k}(t_j), \quad \forall \mathbf{t} = (t_1, \ldots, t_m) \in S_{\mathcal{X}}^m, \text{ and } j = 1, \ldots, m, \\
\mathfrak{M}_n^{(m)}(\varphi, t; \mathbf{h}_{n,k}(\mathbf{t})) &= \widehat{r}_n^{*(m)}(\varphi, t; \mathbf{h}_{n,k}(\mathbf{t})), \\
\mathfrak{M}(\varphi, t) &= r^{(m)}(\varphi, t).
\end{aligned}
$$

Choosing $D_{n,k}^{-}(\xi_n, t)$ and $D_{n,k}^{+}(\xi_n, t)$ such that for all $j = 1, \ldots, m$,

$$
\phi_{t_j}(D_{n,k}^{-}(\xi_n, t_j)) = \frac{\sqrt{\xi_{n,j}} k}{n},
$$

$$\phi_{t_j}(D^+_{n,k}(\xi_n, t_j)) = \frac{k}{n\sqrt{\xi_{n,j}}},$$

where $\xi_{n,j}$ are increasing sequences that belong to $(0, 1)$, and

$$\xi_n = \prod_{j=1}^{m} \xi_{n,j}.$$

We denote $\mathbf{h}^-(\mathbf{t}) = \left(h_1^-(t_1), \ldots, h_m^-(t_m)\right)$ and $\mathbf{h}^+(\mathbf{t}) = \left(h_1^+(t_1), \ldots, h_m^+(t_m)\right)$, where

$$h_j^-(t_j) = D^-_{n,k}(\xi_n, t_j) \text{ and } h^+(t_j) = D^+_{n,k}(\xi_n, t_j), \quad \text{for all } j = 1, \ldots, m.$$

We can easily see that, for all $j = 1, \ldots, m$:

$$\phi_{t_j}^{-1}\left(\frac{\sqrt{\xi_{n,j}}k_{1,n}}{n}\right) \le h^-(t_j) \le \phi_{t_j}^{-1}\left(\frac{\sqrt{\xi_{n,j}}k_{2,n}}{n}\right), \tag{8.47}$$

$$\phi_{t_j}^{-1}\left(\frac{k_{1,n}}{n\sqrt{\xi_{n,j}}}\right) \le h^+(t_j) \le \phi_{t_j}^{-1}\left(\frac{k_{2,n}}{n\sqrt{\xi_{n,j}}}\right). \tag{8.48}$$

Using the condition (2.12) one gets, for all $j = 1, \ldots, m$, there exist constants $0 < \mu_j \le \nu_j < \infty$, such that

$$\mu_j \phi^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right) \le \phi_{t_j}^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right) \text{ and } \phi_{t_j}^{-1}\left(\frac{k_{2,n}}{\rho_{n,j}n}\right) \le \nu_j \phi^{-1}\left(\frac{k_{2,n}}{\rho_{n,j}n}\right),$$

we put $\rho_{n,j} = \sqrt{\xi_{n,j}}$, $h_{n,j} = \mu_j \phi^{-1}\left(\frac{\rho_{n,j}k_{1,n}}{n}\right)$ and $h'_{n,j} = \nu_j \phi^{-1}\left(\frac{k_{2,n}}{\rho_{n,j}n}\right)$, thus $\mathbf{h}^-(\mathbf{t})$ and $\mathbf{h}^+(\mathbf{t})$ belong to the interval

$$\mathcal{H}_n^{(m)} := \prod_{j=1}^{m} \left(h_{n,j}, h'_{n,j}\right).$$

We denote $\widetilde{h}_n = \min_{1 \le j \le m} h_{n,j}$ and $\widetilde{h'}_n = \max_{1 \le j \le m} h'_{n,j}$, therefore,

$$h_j(t_j) \in \left(\widetilde{h}_n, \widetilde{h'}_n\right), \forall j = 1, \ldots, m.$$

We also note for all $\mathbf{b} = (b_1, \ldots, b_m) \in (0, 1)^m$:

$$\mathcal{H}_0^{(m)} := \prod_{j=1}^{m} \left(h_{n,j}, b_j\right),$$

and

$$\widetilde{b}_0 := \max_{1 \le j \le m} b_j.$$

Finally, we can choose constants $0 < \mu^* < \nu^* < \infty$ and a sequence $\{\rho_n^*\} \in (0, 1)$, while respecting the condition **(C.8.)**, in a way that makes $\widetilde{h}_n = \mu^* \phi^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)$, $\widetilde{h'}_n = \nu^* \phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)$ and

$$\mathcal{U}_n = \phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)^{\gamma} + \sqrt{\frac{\psi_{S_X}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^* \phi^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}.$$

It is clear that (L_0) is satisfied due to the condition **(C.3.1)**, and from (8.47) and (8.48), we can easily verify that the construction of $\mathbf{h}^-(\mathbf{t})$ and $\mathbf{h}^+(\mathbf{t})$ satisfies the condition (L_1).

Checking the conditions (L_4) and (L_5)

A direct application of Corollary 3.12 gives

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{\mathbf{h}^-(\mathbf{t})\in\mathcal{H}_n^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\mathfrak{M}_n(\varphi,\mathbf{t};\mathbf{D}_{n,k}^-(\xi_n,\mathbf{t}))-\mathfrak{M}(\varphi,\mathbf{t})\right|$$
$$=O\left(\widetilde{h'}_n^{\gamma}\right)+O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right),$$

which is equivalent to

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{k_{1,n}\le k\le k_{2,n}}\ \sup_{t\in S_{\mathcal{X}}^m}\left|\mathfrak{M}_n(\varphi,\mathbf{t};\mathbf{D}_{n,k}^-(\xi_n,\mathbf{t}))-\mathfrak{M}(\varphi,\mathbf{t})\right|$$
$$=O_{a.co}\left\{\phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)^{\gamma}+\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi_t^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}\right\}. \tag{8.49}$$

Applying the same reasoning with $\mathbf{h}^+(\mathbf{t})$, we obtain

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{k_{1,n}\le k\le k_{2,n}}\ \sup_{t\in S_{\mathcal{X}}^m}\left|\mathfrak{M}_n(\varphi,\mathbf{t};\mathbf{D}_{n,k}^+(\xi_n,\mathbf{t}))-\mathfrak{M}(\varphi,\mathbf{t})\right|$$
$$=O_{a.co}\left\{\phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)^{\gamma}+\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi_t^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}\right\}. \tag{8.50}$$

Thus, the conditions (L_4) and (L_5) are checked.

Checking the condition (L_2)

To check (L_2) we show that for all $\varepsilon_0>0$ and $\mathbf{t}=(t_1,\ldots,t_m)\in S_{\mathcal{X}}^m$:

$$\sum_{n\ge1}\mathbb{P}\left\{\left|\prod_{j=1}^m\mathbb{1}_{\left\{D_{n,k}^-\left(\xi_n,t_j\right)\le H_{n,k}(t_j)\le D_{n,k}^+\left(\xi_n,t_j\right)\right\}}-1\right|>\varepsilon_0\right\}<\infty. \tag{8.51}$$

We have

$$\mathbb{P}\left\{\left|\prod_{j=1}^m\mathbb{1}_{\left\{D_{n,k}^-\left(\xi_n,t_j\right)\le H_{n,k}(t_j)\le D_{n,k}^+\left(\xi_n,t_j\right)\right\}}-1\right|>\varepsilon_0\right\}$$

$$
\begin{aligned}
&= \sum_{j=1}^{m} \mathbb{P}\left\{H_{n,k}(t_j) \notin \left(D_{n,k}^{-}\left(\xi_n, t_j\right), D_{n,k}^{+}\left(\xi_n, t_j\right)\right)\right\} \\
&\leq \sum_{j=1}^{m} \mathbb{P}\left(H_{n,k}(t_j) \leq \phi_{t_j}^{-1}\left(\frac{\sqrt{\xi_{n,j}}k_{1,n}}{n}\right)\right) + \sum_{j=1}^{m} \mathbb{P}\left(H_{n,k}(t_j) \geq \phi_{t_j}^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_{n,j}}n}\right)\right) \\
&\leq \sum_{\ell=1}^{N_\varepsilon(S_\mathcal{X})} \sum_{k=k_{1,n}}^{k_{2,n}} \sum_{j=1}^{m} \mathbb{P}\left(H_{n,k}(t_{\ell(t_j)}) \leq \phi_{t_{\ell(t_j)}}^{-1}\left(\frac{\sqrt{\xi_{n,j}}k_{1,n}}{n}\right)\right) \\
&\quad + \sum_{\ell=1}^{N_\varepsilon(S_\mathcal{X})} \sum_{k=k_{1,n}}^{k_{2,n}} \sum_{j=1}^{m} \mathbb{P}\left(H_{n,k}(t_{\ell(t_j)}) \geq \phi_{t_{\ell(t_j)}}^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_{n,j}}n}\right)\right),
\end{aligned}
$$

now, using

$$
\mathbb{P}\left\{H_{n,k}(t_{\ell(t_j)}) \leq \phi_{t_{\ell(t_j)}}^{-1}\left(\frac{\alpha k_{1,n}}{n}\right)\right\} \leq \exp\left\{-\frac{k - \alpha k_{1,n}}{4}\right\}, \tag{8.52}
$$

and

$$
\mathbb{P}\left\{H_{n,k}(t_{\ell(t_j)}) \geq \phi_{t_{\ell(t_j)}}^{-1}\left(\frac{k_{2,n}}{\alpha n}\right)\right\} \leq \exp\left\{-\frac{(k_{2,n} - \alpha k)^2}{2\alpha k_{2,n}}\right\}. \tag{8.53}
$$

Consequently, we obtain

$$
\begin{aligned}
&\sum_{\ell=1}^{N_\varepsilon(S_\mathcal{X})} \sum_{k=k_{1,n}}^{k_{2,n}} \sum_{j=1}^{m} \mathbb{P}\left\{H_{n,k}(t_{\ell(t_j)}) \leq \phi_{t_{\ell(t_j)}}^{-1}\left(\frac{\alpha k_{1,n}}{n}\right)\right\} \\
&\leq m N_{\varepsilon_n}(S_\mathcal{X}) k_{2,n} \exp\left\{-(1-\alpha)k_{1,n}/4\right\} \\
&\leq m N_{\varepsilon_n}(S_\mathcal{X}) n^{1-\{(1-\alpha)/4\}\frac{k_{1,n}}{\ln n}} < \infty,
\end{aligned} \tag{8.54}
$$

and

$$
\begin{aligned}
&\sum_{\ell=1}^{N_\varepsilon(S_\mathcal{X})} \sum_{k=k_{1,n}}^{k_{2,n}} \sum_{j=1}^{m} \mathbb{P}\left(H_{n,k}(t_{\ell(t_j)}) \geq \phi_{t_{\ell(t_j)}}^{-1}\left(\frac{k_{2,n}}{\sqrt{\xi_{n,j}}n}\right)\right) \\
&\leq N_{\varepsilon_n}(S_\mathcal{X}) k_{2,n} \exp\left\{-(1-\alpha)k_{1,n}/2\alpha\right\} \\
&\leq N_{\varepsilon_n}(S_\mathcal{X}) n^{1-\{(1-\alpha)/2\alpha\}\frac{k_{2,n}}{\ln n}} < \infty.
\end{aligned} \tag{8.55}
$$

Thus (L_2) is checked.

Checking the condition (L_3)

Notice that

$$
\begin{aligned}
&\left|\mathfrak{M}_{n-}^{(m)}(\varphi, \mathbf{t}; \xi_n) - \xi_n \mathfrak{M}^{(m)}(\varphi, \mathbf{t})\right| \\
&= \left| \frac{\sum_{i \in I(m,n)} \varphi(Y_{i_1}, \ldots, Y_{i_m}) \prod_{j=1}^{m} K\left(\frac{d(X_{i_j}, t_j)}{h_j^{-}(t_j)}\right)}{\sum_{i \in I(m,n)} \prod_{j=1}^{m} K\left(\frac{d(X_{i_j}, t_j)}{h_j^{+}(t_j)}\right)} - \frac{\phi_{\mathbf{t}}(\mathbf{h}^{-}(\mathbf{t}))}{\phi_{\mathbf{t}}(\mathbf{h}^{+}(\mathbf{t}))} r^{(m)}(\varphi, \mathbf{t}) \right|
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\phi_{\mathbf{t}}(\mathbf{h}^-(\mathbf{t}))}{\phi_{\mathbf{t}}(\mathbf{h}^+(\mathbf{t}))}\left|\frac{\phi_{\mathbf{t}}(\mathbf{h}^+(\mathbf{t}))}{\phi_{\mathbf{t}}(\mathbf{h}^-(\mathbf{t}))}\frac{\sum_{i\in I(m,n)}\varphi(Y_{i_1},\ldots,Y_{i_m})\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^-(t_j)}\right)}{\sum_{i\in I(m,n)}\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^+(t_j)}\right)} - r^{(m)}(\varphi,\mathbf{t})\right| \\
&= \xi_n\left|\frac{u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))}{u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))} - r^{(m)}(\varphi,\mathbf{t})\right|. \qquad (8.56)
\end{aligned}
$$

The study of (8.56) is similar to the proofs of Theorems 3.10 and 3.11, as we can clearly see that

$$
\begin{aligned}
\left|\frac{u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))}{u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))} - r^{(m)}(\varphi,\mathbf{t})\right| &\leq \left|\frac{u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))}{u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))} - \frac{\mathbb{E}[u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))]}{\mathbb{E}[u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))]}\right| \qquad (8.57) \\
&\quad + \left|\frac{\mathbb{E}[u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))]}{\mathbb{E}[u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))]} - r^{(m)}(\varphi,\mathbf{t})\right|. \qquad (8.58)
\end{aligned}
$$

Let us start with (8.57), we have

$$
\begin{aligned}
&\left|\frac{u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))}{u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))} - \frac{\mathbb{E}[u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))]}{\mathbb{E}[u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))]}\right| \\
&\leq \frac{|u_n(\varphi,\mathbf{t},\mathbf{h}^-(\mathbf{t})) - \mathbb{E}\,(u_n(\varphi,\mathbf{t},\mathbf{h}^-(\mathbf{t})))|}{|u_n(1,\mathbf{t},\mathbf{h}^+)|} \\
&\quad + \frac{|\mathbb{E}\,(u_n(\varphi,\mathbf{t},\mathbf{h}^-(\mathbf{t})))|\cdot|u_n(1,\mathbf{t},\mathbf{h}^+) - \mathbb{E}\,(u_n(1,\mathbf{t},\mathbf{h}^+(\mathbf{t})))|}{|u_n(1,\mathbf{t},\mathbf{h}^+(\mathbf{t}))|\cdot|\mathbb{E}\,(u_n(1,\mathbf{t},\mathbf{h}^+(\mathbf{t})))|} \\
&=: \mathbf{I} + \mathbf{II}.
\end{aligned}
$$

Applying the same calculation as in the proof of Theorem 3.10, it follows that:

$$
\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{(\mathbf{h}^-(\mathbf{t}),\mathbf{h}^+(\mathbf{t}))\in\mathcal{H}_0^{(m)}\times\mathcal{H}_0^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\frac{u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))}{u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))} - r^{(m)}(\varphi,\mathbf{t})\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right).
$$

Now, for (8.58), using the fact that

$$
\mathbb{E}[u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))] \geq \kappa_1^m C_1,
$$

therefore,

$$
\left|\frac{u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))}{u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))} - r^{(m)}(\varphi,\mathbf{t})\right| \leq \kappa_1^{-m}C_1^{-1}\left|u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t})) - r^{(m)}(\varphi,\mathbf{t})\right|,
$$

following the same steps as in Theorem 3.11, we can easily conclude that

$$
\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{(\mathbf{h}^-(\mathbf{t}),\mathbf{h}^+(\mathbf{t}))\in\mathcal{H}_n^{(m)}\times\mathcal{H}_n^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\frac{\mathbb{E}[u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t})]}{\mathbb{E}[u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))]} - r^{(m)}(\varphi,\mathbf{t})\right| = O\left(\widetilde{h'}_n^{\gamma}\right).
$$

Consequently, we have

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{(\mathbf{h}^-(\mathbf{t}),\mathbf{h}^+(\mathbf{t}))\in\mathcal{H}_n^{(m)}\times\mathcal{H}_n^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\frac{u_n(\varphi,\mathbf{t};\mathbf{h}^-(\mathbf{t}))}{u_n(1,\mathbf{t};\mathbf{h}^+(\mathbf{t}))}-r^{(m)}(\varphi,\mathbf{t})\right|$$
$$=O\left(\widetilde{h'}_n^{\gamma}\right)+O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right).$$

Since $\xi_n\to 1$, we can also conclude that

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{(\mathbf{h}^-(\mathbf{t}),\mathbf{h}^+(\mathbf{t}))\in\mathcal{H}_0^{(m)}\times\mathcal{H}_0^{(m)}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\frac{\sum_{i\in I(m,n)}\varphi(Y_{i_1},\ldots,Y_{i_m})\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^-(t_j)}\right)}{\sum_{i\in I(m,n)}\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^+(t_j)}\right)}-\xi_n r^{(m)}(\varphi,\mathbf{t})\right|$$
$$=O\left(\widetilde{h'}_n^{\gamma}\right)+O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(\widetilde{h}_n)}}\right),$$

which implies that

$$\sup_{\varphi\widetilde{K}\in\mathscr{F}_m\mathscr{K}^m}\ \sup_{k_{1,n}\le k\le k_{2,n}}\ \sup_{t\in S_{\mathcal{X}}^m}\left|\frac{\sum_{i\in I(m,n)}\varphi(Y_{i_1},\ldots,Y_{i_m})\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^-(t_j)}\right)}{\sum_{i\in I(m,n)}\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^+(t_j)}\right)}-\xi_n r^{(m)}(\varphi,\mathbf{t})\right|$$
$$=O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)^{\gamma}\right)+O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi_t^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}\right).\tag{8.59}$$

Finally, by putting $\varphi\equiv 1$ in (8.59), we get

$$\sup_{\widetilde{K}\in\mathscr{K}^m}\ \sup_{k_{1,n}\le k\le k_{2,n}}\ \sup_{\mathbf{t}\in S_{\mathcal{X}}^m}\left|\frac{\sum_{i\in I(m,n)}\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^-(t_j)}\right)}{\sum_{i\in I(m,n)}\prod_{j=1}^m K\left(\frac{d(X_{i_j},t_j)}{h_j^+(t_j)}\right)}-\xi_n\right|$$
$$=O\left(\phi^{-1}\left(\frac{k_{2,n}}{\rho_n^* n}\right)^{\gamma}\right)+O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi\left(\mu^*\phi_t^{-1}\left(\frac{\rho_n^* k_{1,n}}{n}\right)\right)}}\right).\tag{8.60}$$

Hence, (L_3) is checked. Now, with all the conditions of Lemma 8.1.1 satisfied, it follows that

$$\sup_{\varphi\widetilde{G}\in\mathscr{F}_m\mathscr{G}^m}\ \sup_{\mathbf{t}\in S_\Omega^m}\ \sup_{k_{1,n}\le k\le k_{2,n}}\left|\mathfrak{M}_n^{(m)}(\varphi,\mathbf{t};\mathbf{D}_{n,k}(\mathbf{t}))-\mathfrak{M}^{(m)}(\varphi,\mathbf{t})\right|=O_{a.co}(\mathcal{U}_n),$$

which is exactly the desired result, hence, the proof is completed.

8.2. Proofs of weak convergence results

Preliminaries of the proofs

As mentioned before, a straightforward approach does not work when dealing with random bandwidths. Therefore, we often use some general lemmas (see, for example, [44]) to be able to use the results of the non-random bandwidths. In this section, we present the results of [147] and [43] obtained for some positive bandwidth h_K. These results are key instrumental in the proofs. We denote the bias term and the centered variate respectively the following quantities

$$\begin{aligned} B_n(t;h_K) &:= \frac{\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t;h_K)\right)-r^{(1)}(\varphi,t)\mathbb{E}\left(\widehat{r}_{n,1}^{(1)}(1,t;h_K)\right)}{\mathbb{E}\left(\widehat{r}_{n,1}^{(1)}(1,t;h_K)\right)},\\ Q_n(t;h_K) &:= \left(\widehat{r}_{n,2}^{(1)}(\varphi,t;h_K)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t;h_K)\right)\right)\\ &\quad -r^{(1)}(\varphi,t)\left(\widehat{r}_{n,1}^{(1)}(1,t;h_K)-\mathbb{E}\left(\widehat{r}_{n,1}^{(1)}(1,t;h_K)\right)\right). \end{aligned} \tag{8.61}$$

The decomposition (8.61) plays a key role in our proof. Indeed, following the method adopted by [147], we will show that convergences in quadratic mean to 1, and that the bias satisfies

$$B_n(t;h_K)=o(1)\text{ as }n\to\infty.$$

Lemma 8.5. *Under conditions **(C.1.)**, **(C.3.1)**, and **(C.5.1)** (or equivalently **(C.1'.)**, **(C.3'.)** and **(C.5.1)**), and if $n\phi(h_K)\to\infty$ as $n\to\infty$, then, we have for each $t\in\mathcal{X}$:*

$$\widehat{r}_{n,1}^{(1)}(1,t;h_K)\xrightarrow{p.s.}1. \tag{8.62}$$

Before we present the next result, the following notation is needed :

$$\mu_{n,1}(t;h_K) := \mathbb{E}\left[\left(\varphi(Y_1)-r^{(1)}(\varphi,t)\right)\Delta_1(t;h_K)\right], \tag{8.63}$$

$$Z_{n,i}(h_K) := \left(\varphi(Y_i)-r^{(1)}(\varphi,t;h_K)\right)\Delta_i(t;h_K)-\mu_{n,1}(t;h_K). \tag{8.64}$$

Set

$$Q_n(t;h_K)=\frac{1}{n\mathbb{E}\left[\Delta_1(t;h_K)\right]}\sum_{i=1}^n Z_{n,i}(h_K),$$

and

$$\sigma_{n,0}^2(t;h_K):=\frac{1}{\mathbb{E}^2\left[\Delta_1(t;h_K)\right]}\mathrm{Var}\left[Z_{n,1}(h_K)\right].$$

Lemma 8.6. *Under conditions **(C.1.)**, **(C.2.)**, **(C.3.1)** and **(C.5.1)** (with $p=\infty$ and $\delta>1$ for condition **(C.5.1)**), we have for $n\phi(h_K)\underset{n\to\infty}{\longrightarrow}\infty$, $t\in\mathcal{X}$ and positive constants C_4,C_5:*

$$C_4 \frac{g_2(t,\varphi)}{f_1(t)} \leq \phi(h_K)\sigma^2_{n,0}(t) \leq C_5 \frac{g_2(t,\varphi)}{f_1(t)}, \tag{8.65}$$

whenever $f_1(t) > 0$. We have

$$\frac{1}{n\mathbb{E}^2\left[\Delta_1(t;h_K)\right]} \sum_{i=1}^{n} \sum_{\substack{i=1 \\ i\neq j}}^{n} \mathrm{Cov}\left[Z_{n,i}, Z_{n,j}\right] = o\left(\sigma^2_{n,0}(t)\right). \tag{8.66}$$

$$\mathrm{Var}\left[Q_n(t)\right] = \sigma^2_{n,0}(t)(1+o(1))/n. \tag{8.67}$$

Lemma 8.7. *Under conditions* ***(C.1'.)***, ***(C.2.)***, ***(C.3'.)*** *and* ***(C.5.1)*** *(with $p = \infty$ and $\delta > 1$ for condition* ***(C.5.1)****), we have for $n\phi(h_K) \to \infty$ as $n \to \infty$ and $t \in \mathcal{X}$:*

$$\phi(h_K)\sigma^2_{n,0}(t) \longrightarrow \frac{\kappa'_1}{\kappa'_2} \frac{g_2(t,\varphi)}{f_1(t)}, \tag{8.68}$$

whenever $f_1(t) > 0$, and κ'_1, κ'_2 are constants specified previously.

$$\frac{1}{n\mathbb{E}^2\left[\Delta_1(t;h_K)\right]} \sum_{i=1}^{n} \sum_{\substack{i=1 \\ i\neq j}}^{n} \mathrm{Cov}\left[Z_{n,i}, Z_{n,j}\right] = o\left(\sigma^2_{n,0}(t)\right), \tag{8.69}$$

$$n\phi(h_K)\mathrm{Var}\left[Q_n(t)\right] \longrightarrow \frac{\kappa'_1}{\kappa'_2} \frac{g_2(t,\varphi)}{f_1(t)}. \tag{8.70}$$

To unburden the notation a bit and for simplicity, we denote

$$\sigma^2(t) := \left(\kappa'_1 g_2(t,\varphi)\right)/\left(\kappa'_2 f_1(t)\right).$$

Lemma 8.8. *Under conditions* ***(C.1'.)***, ***(C.2.)***, ***(C.3'.)***, ***(C.5.1)*** *and* ***(C.5.2)***, *if $n\phi(h_K) \to \infty$ as $n \to \infty$, then we have for $t \in \mathcal{X}$, as $n \to \infty$,*

$$(n\phi(h_K))^{1/2} Q_n(t) \longrightarrow \mathcal{N}\left(0, \sigma^2(t)\right). \tag{8.71}$$

Lemma 8.9. *Under conditions* ***(C.1.)***, ***(C.2.)***, ***(C.3.1)*** *and* ***(C.5.1)***, *and if*

$$\left(n\phi(h_K)/\log\log n\right) \underset{n\to\infty}{\longrightarrow} \infty,$$

then we have, as $n \to \infty$,

$$\left(\frac{n\phi(h_K)}{\log\log n}\right)^{1/2} \left[\widehat{r}^{(1)}_n(\varphi,t) - r^{(1)}(\varphi,t) - B_n(t)\right] \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{8.72}$$

Proof of Theorem 4.1

Using the Cramér-Wold device, it is sufficient to prove the convergence of one-dimensional distribution in order to prove Theorem 4.1. Indeed, by the linearity of $\nu_n(\psi; k_K \mid t)$ if suffices to show that

$$\nu_n(\Phi; H_{n,k}(t) \mid t) \rightarrow N(0, \sigma^2(\Phi, t)),$$

for all Φ of the form

$$\Phi = \sum_{p=1}^{L} c_p \boldsymbol{\psi}_p, \quad c_1, \ldots, c_L \in \mathbb{R}, \quad \boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_L \in \mathscr{F}\mathscr{K}.$$

Therefore, we shall only demonstrate convergence in a single dimension. Remember that we're dealing with

$$\begin{aligned} \nu_n(\psi; H_{n,k}(t) \mid t) &= \sqrt{k}(\widehat{r}_n^{(1)}(\psi, t; H_{n,k}(t)) - r^{(1)}(\varphi, t)) \\ &= \sqrt{k}\left(\frac{\sum_{i=1}^n \varphi(Y_i) K\left(d(X_i, t)/H_{n,k}(t)\right)}{\sum_{i=1}^n K\left(d(X_i, t)/H_{n,k}(t)\right)} - r^{(1)}(\varphi, t)\right). \end{aligned} \tag{8.73}$$

Set

$$\varphi_M(Y) := \varphi(Y) \mathbb{1}_{\{F(Y) < M\}}.$$

To obtain the desired result, we write

$$\begin{aligned} &\widehat{r}_n^{(1)}(\varphi, t; H_{n,k}(t)) - r^{(1)}(\varphi, t) \\ &= \frac{\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t))}{\widehat{r}_{n,1}^{(1)}(1, t; H_{n,k}(t))} - r^{(1)}(\varphi, t) \\ &= \frac{1}{\widehat{r}_{n,1}^{(1)}(1, t; H_{n,k}(t))}\left[\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t)) - \widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t))\right)\right. \\ &\quad \left. - \left(r^{(1)}(\varphi, t) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t))\right)\right)\right] \\ &\quad + \left(\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t))\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t))\right)\right) \\ &\quad + \left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t)) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t))\right) - r^{(1)}(\varphi, t)\left(\widehat{r}_{n,1}^{(1)}(1, t; H_{n,k}(t)) - 1\right)\right) \\ &= \frac{1}{\widehat{r}_{n,1}^{(1)}(1, t; H_{n,k}(t))}\left[I_1(t; H_{n,k}) + I_2(t; H_{n,k})\right], \end{aligned}$$

where

$$\begin{aligned} I_1(t; H_{n,k}) &= \underbrace{\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t)) - \widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t))\right)}_{S_1(t; H_{n,k})} - \mathbb{E}\left(S_1(t; H_{n,k})\right) \\ &\quad + \underbrace{\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t))\right) - r^{(1)}(\varphi, t)}_{S_2(t; H_{n,k})}, \end{aligned} \tag{8.74}$$

and

$$I_2(t;H_{n,k}) = \left(\widehat{r}^{(1)}_{n,2}(\varphi_M,t;H_{n,k}(t)) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M,t;H_{n,k}(t))\right) - r^{(1)}(\varphi,t)\left(\widehat{r}^{(1)}_{n,1}(1,t;H_{n,k}(t)) - 1\right)\right). \tag{8.75}$$

To obtain the desired results, we follow the strategy of [152].

Lemma 8.10. *Under the assumptions **(C.2.1), (C.2.2), (C.3.2), (C.4.2)** and **(C.8.)**, we have*

$$\sqrt{k} I_1\left(t, H_{n,k}(t)\right) \xrightarrow{\mathbb{P}} 0, \quad as \quad n\to\infty$$

Lemma 8.11. *Under the assumptions **(C.1'.), (C.2.1.), (C.2.2.), (C.3'.)** and **(C.5.)** and if*

$$k\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^{2\gamma} \to 0 \text{ as } n\to\infty,$$

we have

$$\sqrt{k} I_2(\varphi,t,H_{n,k}(t)) \to \mathcal{N}\left(0,\sigma^2(t)\right), \quad \text{as } n\to\infty.$$

Lemma 8.12. *Under the assumptions **(C.1), (C.3.1),** and **(C.5.1)** (or equivalently **(C.1'.), (C.3'.)** and **(C.5.1)**), and if*

$$k\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^{2\gamma} \to 0 \text{ as } n\to\infty,$$

then we have for each $t\in\mathcal{X}$:

$$\widehat{r}^{(1)}_{n,1}(1,t,H_{n,k}(t)) - 1 \xrightarrow{\mathbb{P}} 0 \text{ as } n\to\infty.$$

Lemma 8.13. *Under conditions **(C.1.), (C.3.1),** and **(C.5.1)** (or equivalently **(C.1'.), (C.3'.)** and **(C.5.1)**), and if $n\phi(h_K)\to\infty$ as $n\to\infty$. Then we have for each $t\in\mathcal{X}$:*

$$\widehat{r}^{(1)}_{n,1}(\varphi,t;h_K) \xrightarrow{p.s.} \mathbb{E}(\widehat{r}^{(1)}_{n,1}(\varphi,t;h_K)), \tag{8.76}$$

as $n\to\infty$.

We highlight that this lemma is more general than Lemma 8.5. This result is slightly weaker than the uniform, almost complete convergence with the rate that we obtained in Section 3.14. However, the conditions imposed in this lemma are less restrictive.

Proof of Theorem 4.2

In this section, we will use the same method as in [43] and earlier [8], that is using the blocking approach which entails breaking down a strictly stationary sequence $(X_1,\ldots,X_n)$, into $2\upsilon_n$, equal-sized blocks, that each one is of length $n-2\upsilon_n a_n$, keeping in mind the notation given in the proof of Lemma 8.2. In order to establish the asymptotic equi-continuity of the conditional empirical process

$$\left\{\nu_n(\psi\mid t) = \sqrt{k}\left(\widehat{r}^{(1)}_n(\psi,t;H_{n,k}(t)) - r^{(1)}(\varphi,t)\right), \psi\in\mathscr{F}\mathscr{K}, \varphi\in\mathscr{F}\right\}.$$

Let us introduce, for any $\varphi K \in \mathscr{F}\mathscr{K}$ and $t \in \mathcal{X}$,

$$\begin{aligned} \mathcal{W}_n(t,\varphi;H_{n,k}(t)) &:= \sum_{i=1}^{n} \varphi(Y_i) K\left(\frac{d(X_i,t)}{H_{n,k}(t)}\right) - n\mathbb{E}\left\{\varphi(Y_1)K\left(\frac{d(X_1,t)}{H_{n,k}(t)}\right)\right\}, \quad (8.77)\\ \nu_n(\varphi \mid t) &= \sqrt{k}\left(\widehat{r}_n^{(1)}(\varphi,t;H_{n,k}(t)) - r^{(1)}(\varphi,t)\right)\\ &:= \sqrt{k}\left(\widehat{r}_n^{(1)}(\psi,t,H_{n,k}(t)) - r^{(1)}(\varphi,t)\right). \quad (8.78)\end{aligned}$$

Then, we have

$$\begin{aligned} \nu_n(\psi \mid t) &= \sqrt{k}(\widehat{r}_n^{(1)}(\psi,t,H_{n,k}(t)) - r^{(1)}(\varphi,t))\\ &= \sqrt{k}\left(\frac{\sum_{i=1}^{n}\varphi(Y_i)K\left(d(X_i,t)/H_{n,k}(t)\right)}{\sum_{i=1}^{n}K\left(d(X_i,t)/H_{n,k}(t)\right)} - r^{(1)}(\varphi,t)\right)\\ &= \frac{1}{\breve{r}_{n,1}^{(1)}(\varphi,t;H_{n,k}(t))}\frac{1}{\sqrt{k}}\mathcal{W}_n(t,\varphi;H_{n,k}(t)) - \frac{\mathbb{E}(\breve{r}_{n,2}^{(1)}(\varphi,t;H_{n,k}(t)))}{\breve{r}_{n,1}^{(1)}(\varphi,t)\mathbb{E}(\breve{r}_{n,1}^{(1)}(\varphi,t;H_{n,k}(t)))}\frac{1}{\sqrt{k}}\mathcal{W}_n(t,1;H_{n,k}(t))\\ &\quad - \sqrt{k}B_n(t;H_{n,k}(t)),\end{aligned}$$

where for $h_K > 0$, we have

$$\begin{aligned} \breve{r}_{n,2}^{(1)}(\varphi,t,h_K) &:= \frac{1}{n\phi(h_K(t))}\sum_{i=1}^{n}\varphi(Y_i)\Delta_i(t,h_K),\\ \breve{r}_{n,1}^{(1)}(1,t,h_K) &:= \frac{1}{n\phi(h_K(t))}\sum_{i=1}^{n}\Delta_i(t,h_K).\end{aligned}$$

We study the asymptotic equi-continuity of each of the previous terms. For a class of functions $\mathscr{G}$, let $\alpha_n(\cdot)$ be an empirical process based on $(X_1,Y_1),\ldots,(X_n,Y_n)$ and indexed by $\mathscr{G}$:

$$\alpha_n(g) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{g(X_i,Y_i) - \mathbb{E}\left(g(X_i,Y_i)\right)\right\}, \quad \text{with} \quad \|\alpha_n(g)\|_{\mathscr{G}} = \sup_{g\in\mathscr{G}}|\alpha_n(g)|,$$

and for a measurable function $\varphi(\cdot)$ and $t \in \mathcal{X}$, set

$$\eta_{n,t,\varphi,K}(u,v,h_K) = \varphi(v)K\left(\frac{d(u,t)}{h_K}\right), \quad \text{for } u,v \in \mathcal{X}.$$

That implies

$$\frac{1}{\sqrt{n\phi(h_K)}}\mathcal{W}_n(t,\varphi,h_K) = \frac{1}{\sqrt{\phi(h_K)}}\alpha_n(\eta_{n,t,\varphi,K}).$$

Again, keeping in mind that $\mathbb{1}_{\left\{D_n^- \leq H_{n,k}(t) \leq D_n^+\right\}} \stackrel{a.co}{\longrightarrow} 1$ when $\frac{k}{n} \to 0$. We will establish the asymptotic equi-continuity of

$$\left\{\sqrt{\frac{n}{k}}\alpha_n(\eta_{n,t,\varphi,K}) : \eta_{n,t,\varphi,K} \in \mathscr{F}\mathscr{K}\right\},$$

which means, for every $\varepsilon > 0$, that

$$\lim_{b\to 0}\limsup_{n\to\infty}\mathbb{P}\left\{\sqrt{\frac{n}{k}}\left\|\alpha_n\left(\eta_{n,t,\varphi,K}\right)\right\|_{\mathscr{FK}(b,\|\cdot\|_p)} > \varepsilon\right\} = 0,$$

where

$$\mathscr{FK}(b,\|\cdot\|_p) = \left\{\eta_{n,t,\varphi_1,K_1} - \eta_{n,t,\varphi_2,K_2} : \left\|\eta_{n,t,\varphi_1,K_1} - \eta_{n,t,\varphi_2,K_2}\right\|_p < b,\ \eta_{n,t,\varphi_1,K_1}, \eta_{n,t,\varphi_2,K_2} \in \mathscr{FK}\right\}.$$

The idea is to work with the independent block sequence $\left\{\xi_j = (\zeta_j, \varsigma_j)\right\}_{j=1}^{\infty}$ instead of working on dependent one, which is possible through (A.1), then we have

$$\begin{aligned}
&\mathbb{P}\left\{\left\|k^{-1/2}\sum_{j=1}^{n}\left(\varphi(Y_j)K\left(\frac{d\left(X_j,t\right)}{H_{n,k}(t)}\right) - \mathbb{P}(\eta_{n,t,\varphi,K}(H_{n,k}(t)))\right)\right\|_{\mathscr{FK}(b,\|\cdot\|_p)} > \delta\right\}\\
&\quad\le 2\mathbb{P}\left\{\left\|k^{-1/2}\sum_{j=1}^{\upsilon_n}\sum_{i\in H_j}\left(\varphi(\zeta_i)K\left(\frac{d\left(\varsigma_i,t\right)}{H_{n,k}(t,\varsigma)}\right) - \mathbb{P}(\eta_{n,t,\varphi,K}(H_{n,k}(t,\varsigma)))\right)\right\|_{\mathscr{FK}(b,\|\cdot\|_p)} > \delta'\right\}\\
&\quad + 2(\upsilon_n - 1)\beta_{a_n},
\end{aligned} \tag{8.79}$$

where $H_{n,k}(t,\varsigma)$ is defined by

$$H_{n,k}(t,\varsigma) = \min\left\{h\in\mathbb{R}^+ : \sum_{i=1}^{n}\mathbb{1}_{B(t,h)}(\varsigma_i) = k\right\}, \tag{8.80}$$

We choose

$$a_n = [(\log n)^{-1}(n^{p-2}\phi^p(h_K))^{1/2(p-1)}] \text{ and } \upsilon_n = \left[\frac{n}{2a_n}\right] - 1.$$

Note that a_n in our setting is equivalent to:

$$a_n = (\log n)^{-1}(n^{-2}k^p)^{1/2(p-1)}.$$

Making use of the condition **(C.5.1)**, we get $(\upsilon_n - 1)\beta_{a_n} \longrightarrow 0$ as $n \to 0$, then it's just a matter of the right side term of (8.79). Let us begin with, the blocks being independent, we symmetrize using a sequence $\{\epsilon_j\}_{j\in\mathbb{N}^*}$ of i.i.d. Rademacher variables, i.e., r.v's with

$$\mathbb{P}(\epsilon_j = 1) = \mathbb{P}(\epsilon_j = -1) = 1/2.$$

It should be noted that the sequence $\{\epsilon_j\}_{j\in\mathbb{N}^*}$ is independent of the sequence $\{\boldsymbol{\xi}_i = (\varsigma_i, \zeta_i)\}_{i\in\mathbb{N}^*}$, thus it remains to establish, for all $\delta > 0$,

$$\lim_{b\to 0}\limsup_{n\to\infty}\mathbb{P}\left\{\left\|k^{-1/2}\sum_{j=1}^{\upsilon_n}\epsilon_j\sum_{i\in H_j}\left(\varphi(\zeta_i)K\left(\frac{d\left(\varsigma_i,t\right)}{H_{n,k}(t,\varsigma)}\right)\right)\right\|_{\mathscr{FK}(b,\|\cdot\|_p)} > \delta\right\} = 0.$$

Again, using the fact that that $\mathbb{1}_{\{D_n^- \leq H_{n,k}(t,\varsigma) \leq D_n^+\}} \xrightarrow{a.co} 1$ when $\frac{k}{n} \to 0$ and (A.9), it suffices to show that

$$\lim_{b\to 0}\limsup_{n\to\infty}\mathbb{P}\left\{\left\|k^{-1/2}\sum_{j=1}^{\upsilon_n}\epsilon_j\sum_{i\in H_j}\left(\varphi(\zeta_i)K\left(\frac{d(\varsigma_i,t)}{D_n^-}\right)\right)\right\|_{\mathscr{F}\mathscr{K}(b,\|\cdot\|_p)} > \delta\right\} = 0.$$

Since the p^{th}-conditional moment satisfies **(C.4.2)**, we can truncate and we get, for each $\lambda > 0$, as $n \to \infty$,

$$\begin{aligned} & k^{-1/2}\sum_{j=1}^{n}\mathbb{E}\left(\kappa_2 F(\zeta_i)\mathbb{1}_{\{F(\zeta_i)\geq\lambda(M_n)^{1/2(p-1)}\}}\right) && (8.81)\\ = & \quad k^{-1/2}\int_0^\infty \mathbb{P}\left(\kappa_2 F\mathbb{1}_{\{F\geq\lambda(M_n)^{1/2(p-1)}\}} \geq t\right)dt \\ = & \quad k^{-1/2}\int_0^{\lambda(M_n)^{1/2(p-1)}}\mathbb{P}\left(F \geq \lambda(M_n)^{1/2(p-1)}\right)dt \\ & +k^{-1/2}\int_{\lambda(M_n)^{1/2(p-1)}}^\infty \mathbb{P}(F\geq t)dt \\ \underset{n\to\infty}{\longrightarrow} & \quad 0. && (8.82)\end{aligned}$$

Hence,

$$\exists\lambda_n \underset{n\to\infty}{\to} 0: \qquad k^{-1/2}\mathbb{E}\left(\kappa_2 F\mathbb{1}_{\{F\geq\lambda_n(M_n)^{1/2(p-1)}\}}\right)\underset{n\to\infty}{\longrightarrow} 0.$$

Then, it suffices to show

$$\lim_{b\to 0}\limsup_{n\to\infty}\mathbb{P}\left\{\left\|k^{-1/2}\sum_{j=1}^{\upsilon_n}\epsilon_j\sum_{i\in H_j}\left(\varphi(\zeta_i)K\left(\frac{d(\boldsymbol{\varsigma}_i,t)}{D_n^-}\right)\right)\mathbb{1}_{\{\kappa_2 F(\zeta_i)\leq\lambda_n M_n^{1/2(p-1)}\}}\right\|_{\mathscr{F}\mathscr{K}(b,\|\cdot\|_p)} > \delta\right\} = 0.$$

We have the following

$$\nu_n^{(2)}(\eta_{n,t,\varphi,K}) = k^{-1/2}\sum_{j=1}^{\upsilon_n}\epsilon_j\sum_{i\in H_j}\left(\varphi(\zeta_i)K\left(\frac{d(\boldsymbol{\varsigma}_i,t)}{D_n^-}\right)\right)\mathbb{1}_{\{\kappa_2 F(\zeta_i)\leq\lambda_n M_n^{1/2(p-1)}\}}.$$

This is done by using the chaining method. [8] gave $b_q = b2^{-q}$, $q = 0,\ldots,q_n$ where q_n

$$2^{-1}\lambda_n(\log(n))^{-1} \leq b_{q_n}^2 \leq \lambda_n(\log(n))^{-1}, \tag{8.83}$$

and let the class $\mathscr{F}\mathscr{K}_q$ of measurable functions of $\mathscr{F}\mathscr{K}$

$$\#\mathscr{F}\mathscr{K}_q = N_q := N(b_q,\mathscr{F}\mathscr{K},\|\cdot\|_p) \qquad \sup_{\eta_{n,t,\varphi_1,K_1}\in\mathscr{F}\mathscr{K}}\ \min_{\eta_{n,t,\varphi_2,K_2}\mathscr{F}\mathscr{K}_q}\left\|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\right\|_p \leq b_q.$$

There is a map $\pi_q : \mathscr{F}\mathscr{K} \longrightarrow \mathscr{F}\mathscr{K}_q$ that takes each $\eta_{n,t,\varphi,K} \in \mathscr{F}\mathscr{K}$ to its closest function in $\mathscr{F}\mathscr{K}_q$ such that

$$\left\|\eta_{n,t,\varphi,K} - \pi_q(\eta_{n,t,\varphi,K})\right\|_p \leq b_q.$$

Applying the chaining method

$$\begin{aligned}
&\sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in\mathscr{F}\mathscr{K}\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le b}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2})\\
&\le \sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in\mathscr{F}\mathscr{K}\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le b q_n}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2})\\
&\quad +2\sum_{q=1}^{q_n}\sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in(\mathscr{F}\mathscr{K})_{q-1}\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le 3b q}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2})\\
&\quad +\sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in(\mathscr{F}\mathscr{K})_0\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le 2b}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}).
\end{aligned}\tag{8.84}$$

Let δ_q be in such a way that

$$\delta_q=(b_q)^{1/2}\vee(3b_q(8+c_{p,\beta}^2)^{1/2}(\log N_q)^{1/2}).\tag{8.85}$$

Let r be chosen so small in such a way that

$$2\sum_{q=1}^{+\infty}\delta_q\le\delta.$$

Therefore, from (8.84), we readily infer that

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in\mathscr{F}\mathscr{K}\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le b}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2})\ge 3\delta\right\}\\
&\le \mathbb{P}\left\{\sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in\mathscr{F}\mathscr{K}\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le b q_n}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2})\ge\delta\right\}\\
&\quad +2\sum_{q=1}^{q_n}\mathbb{P}\left\{\sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in(\mathscr{F}\mathscr{K})_{q-1}\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le 3b q}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2})\ge\delta_q\right\}\\
&\quad +\mathbb{P}\left\{\sup_{\substack{\eta_{n,t,\varphi_1,K_1},\eta_{n,t,\varphi_2,K_2}\in(\mathscr{F}\mathscr{K})_0\\ \|\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2}\|_p\le 2b}} \nu_n^{(2)}(\eta_{n,t,\varphi_1,K_1}-\eta_{n,t,\varphi_2,K_2})\ge\delta\right\}=:\mathbb{A}+\mathbb{B}+\mathbb{C}.
\end{aligned}$$

By the fact that the terms composing $\nu_n^{(2)}(\eta_{n,t,\varphi,K})$ are bounded by $a_n\lambda_n M_n^{1/2(p-1)}$, and by applying the Bernstein's inequality, we obtain

$$\mathbb{B} \leq 2\sum_{q=1}^{q_n} \exp\left(2\log N_q - \frac{\delta_q^2 k}{nb_q^2 c_{p,\beta}^2 + (4/3)\delta_q a_n\lambda_n n^{p/2(p-1)}\phi(h_K)^{(p-2)/2(p-1)}}\right).$$

By using (8.83), we have

$$\delta_q a_n\lambda_n n^{p/2(p-1)}\phi(h_K)^{(p-2)/2(p-1)} = (4/3)\delta_q\lambda_n k(\log(n))^{-1} \leq (8/3)nb_q^2\delta_q \leq 8nb_q^2,$$

that means

$$\begin{aligned}\mathbb{B} &\leq 2\sum_{q=1}^{q_n} \exp\left(2\log N_q - \frac{\delta_q^2}{(8+c_{p,\beta}^2)b_q^2}\right) \leq 2\sum_{q=1}^{q_n} \exp\left(-\frac{\delta_q^2}{2(8+c_{p,\beta}^2)b_q^2}\right)\\ &\leq 2\sum_{q=1}^{\infty} \exp\left(-\frac{2^q}{2(8+c_{p,\beta}^2)b}\right) \longrightarrow 0 \quad as \quad b\to 0. \end{aligned} \tag{8.86}$$

In view of (8.85), we assume that $\delta < 3$. In a similar way, we have

$$\mathbb{C} \leq 2\exp\left(2\log N_0 - \frac{\delta^2}{(8+c_{p,\beta}^2)b^2}\right) \longrightarrow 0 \quad as \quad b\to 0.$$

Finally, by (8.83) it suffices to prove, for each $\delta > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left\{\left\|\nu_n^{(2)}(\eta_{n,t,\varphi,K})\right\|_{\mathscr{F}\mathscr{K}_{(\lambda_n^{1/2}(\log(n))^{-1/2},\|\cdot\|_p)}} \geq \delta\right\} = 0.$$

Making use of the square root trick (Lemma 5.2 [102]) and see also [137] in similar way as in [8], we get

$$\begin{aligned} &\mathbb{P}\left\{\left\|k^{-1/2}\sum_{j=1}^{\upsilon_n}\epsilon_j\sum_{i\in H_j}\left(\varphi(\zeta_i)K\left(\frac{d(\varsigma_i,t)}{D_n^-}\right)\right)\right.\right.\\ &\qquad\left.\left.\times\mathbb{1}_{\{\kappa_2 F(\zeta_i)\leq\lambda_n(n/\phi(h_K))^{1/2(p-1)}\}}\right\|_{\mathscr{F}\mathscr{K}_{(\lambda_n^{1/2}(\log(n))^{-1/2},\|\cdot\|_p)}} \geq 2\delta\right\}\\ &\leq \mathbb{P}\left\{\left\|k^{-1/2}\sum_{j=1}^{\upsilon_n}\epsilon_j\sum_{i\in H_j}\varphi(\zeta_i)K\left(\frac{d(\varsigma_i,t)}{D_n^-}\right)\right.\right.\\ &\qquad\left.\times\mathbb{1}_{\{\kappa_2 F(\zeta_i)\leq\lambda_n(n/\phi(h_K))^{1/2(p-1)}\}}\right\|_{\mathscr{F}\mathscr{K}_{(\lambda_n^{1/2}(\log(n))^{-1/2},\|\cdot\|_p)}} \geq 2\delta,\\ &\qquad\left\|k^{-1}\sum_{j=1}^{\upsilon_n}\left(\sum_{i\in H_j}\varphi(\zeta_i)K\left(\frac{d(\varsigma_i,t)}{D_n^-}\right)\mathbb{1}_{\{\kappa_2 F(\zeta_i)\leq\lambda_n(n/\phi(h_K))^{1/2(p-1)}\}}\right)^2\right\|_{\mathscr{F}\mathscr{K}_{(\lambda_n^{1/2}(\log(n))^{-1/2},\|\cdot\|_p)}}\\ &\qquad\left.\leq 64\lambda_n c_{p,\beta}^2(\log(n))^{-1}\right\} \end{aligned} \tag{8.87}$$

$$+\mathbb{P}\left\{\left\|k^{-1}\sum_{j=1}^{\upsilon_n}\left(\sum_{i\in H_j}\varphi(\zeta_i)K\left(\frac{d(\varsigma_i,t)}{D_n^-}\right)\mathbb{1}_{\{\kappa_2 F(\zeta_i)\leq\lambda_n(n/\phi(h_K))^{1/2(p-1)}\}}\right)^2\right\|_{\mathscr{F}\mathscr{K}_{(\lambda_n^{1/2}(\log(n))^{-1/2},\|\cdot\|_p)}}\right.$$

$$\left. > 64\lambda_n c_{p,\beta}^2(\log(n))^{-1}\right\}$$

$$=: \quad \mathbb{P}(\mathbb{A}_1)+\mathbb{P}(\mathbb{A}_2).$$

Let us introduce the semi-norm

$$\widetilde{d}_{n\phi,2} := \left(k^{-1}\sum_{j=1}^{\upsilon_n}\sum_{i\in H_j}\left|\eta_{n,t,\varphi_1,K_1}(\varsigma_i,\zeta_i)-\eta_{n,t,\varphi_2,K_2}(\varsigma_i,\zeta_i)\right|^2\right)^{1/2},$$

and the covering number defined for any class of functions $\mathscr{E}$ by

$$\widetilde{N}_{n\phi,2}(u,\mathscr{E}) := N_{n\phi,2}(u,\mathscr{E},\widetilde{d}_{n\phi,2}).$$

By the latter we can bound $\mathbb{P}(\mathbb{A}_1)$, (the calculations are detailed in [27]). In the same way, as in [27] and before in [8], as a result of the independence between the blocks and condition **(C.4.3)**, we apply again Lemma 5.2 in [102] and get

$$\mathbb{P}(\mathbb{A}_2)\to 0.$$

Therefore, the theorem is proved.

Proof of Theorem 4.5

Lemma 8.14. *Under assumptions of Theorem 4.5, we have*

$$\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left(\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))-\mathbb{E}\left[u_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))\right]\right)\xrightarrow{d}\mathcal{N}\left(0,m^2\sigma_{\mathbf{t}}^2(\varphi,\varphi)\right), \tag{8.88}$$

and, if in addition, condition **(C.8)** is satisfied, then we have

$$\sqrt{n^{-m+1}k^m}\left\{\widehat{r}_n^{*(m)}(\varphi,\mathbf{t};\mathbf{h}_{n,k}(\mathbf{t}))-\mathbb{E}\left[u_n(\varphi,\mathbf{t};\mathbf{h}(\mathbf{t}))\right]\right\}\xrightarrow{d}\mathcal{N}\left(0,m^2\sigma_{\mathbf{t}}^2(\varphi,\varphi)\right). \tag{8.89}$$

Lemma 8.15. *Under the assumptions of Theorem 4.5, we have*

$$\sqrt{n\tilde{\phi}(\mathbf{h})}\left(u_n(\varphi_1,\mathbf{t},\mathbf{h})-\mathbb{E}\left(u_n(\varphi_1,\mathbf{t},\mathbf{h}\right),u_n(\varphi_2,\mathbf{t},\mathbf{h})-\mathbb{E}\left(u_n(\varphi_2,\mathbf{t},\mathbf{h})\right)\xrightarrow{d}\mathcal{N}(0,\Sigma), \tag{8.90}$$

where

$$\Sigma := \left(m^2\sigma_{\mathbf{t}}^2\left(\varphi_i,\varphi_j\right)\right)_{i,j=1,2}.$$

Proof of Theorem 4.5

As mentioned earlier, the study of the weak convergence of the conditional U-process is based on the study of two parts: the truncated part and the remainder part.

Lemma 8.16. *Let $\mathscr{F}_m\mathscr{K}^m$ be a uniformly bounded class of measurable canonical functions from $\mathcal{X}^m \times \mathcal{Y}^m \longrightarrow \mathbb{R}$, $m \geq 2$. Suppose that there are finite constants $\boldsymbol{a}$ and $\boldsymbol{b}$ such that the $\mathscr{F}_m\mathscr{K}^m$ covering number satisfies:*

$$N(\epsilon, \mathscr{F}_m\mathscr{K}^m, \|\cdot\|_{L_2(Q)}) \leq \mathbf{a}\epsilon^{-\mathbf{b}}, \tag{8.91}$$

for every $\epsilon > 0$ and every probability measure Q. If the mixing coefficient β of the stationary sequence $\{Z_i = (X_i, Y_i)\}_{i\in\mathbb{N}^*}$ fulfills

$$\beta_s s^r \longrightarrow 0, \ as\ s \to \infty \tag{8.92}$$

for some $r > 1$, then

$$\left\| n^{\frac{-3m+1}{2}} \sqrt{k^m} \sum_{\mathbf{i}\in I_m^n} G_{\varphi,\mathbf{t},\mathbf{h}_{n,k}}(\mathbf{X}_i, \mathbf{Y}_i) \right\|_{\mathscr{F}_m\mathscr{K}^m} \xrightarrow{\mathbb{P}} 0.$$

Proof of Theorem 4.7

It is known that the weak convergence of an empirical process is obtained from its finite-dimensional convergence and its asymptotic equi-continuity (while respecting certain criteria). Theorem 4.5 gives the finite-dimensional convergence of the conditional U-process $\{\mu_n(\varphi, \mathbf{t}, \mathbf{h}(\mathbf{t}))\}_{\mathscr{F}_m\mathscr{K}^m}$, so what remains to be seen is its asymptotic equi-continuity. We decompose the U-process $\mu_n(\varphi, \mathbf{t})$ into two parts truncated and remainder part:

$$\mu_n(\varphi, \mathbf{t}, \mathbf{h}(\mathbf{t})) = \mu_n^{(T)}(\varphi, \mathbf{t}, \mathbf{h}(\mathbf{t})) + \mu_n^{(R)}(\varphi, \mathbf{t}, \mathbf{h}(\mathbf{t})).$$

Following the same reasoning to obtain (A.48), we also know that:

$$\mu_n(\varphi, \mathbf{t}; \mathbf{D}_n^+) \leq \mu_n(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(t)) \leq \mu_n(\varphi, \mathbf{t}; \mathbf{D}_n^-). \tag{8.93}$$

Therefore, it suffices to prove the weak convergence of $\mu_n(\varphi, \mathbf{t}; \mathbf{D}_n^-)$ and $\mu_n(\varphi, \mathbf{t}; \mathbf{D}_n^+)$ instead of studying $\mu_n(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(t))$. The steps of the proof are similar to [43] while taking into account a multivariate bandwidth $\mathbf{h}(\mathbf{t})$ instead of a univariate bandwidth h_K. In this section we only show the proof for $\mu_n^{(T)}(\varphi, \mathbf{t}; \mathbf{D}_n^-)$, the process can be replicated for $\mu_n^{(T)}(\varphi, \mathbf{t}; \mathbf{D}_n^+)$.
As shown earlier, the truncated part $\mu_n^{(T)}(\varphi, \mathbf{t}, \mathbf{h}(\mathbf{t}))$ is decomposed according to the Hoeffding's decomposition:

$$\mu_n^{(T)}(\varphi, \mathbf{t}, \mathbf{D}_n^-) = \sqrt{n\tilde{\phi}(\mathbf{D}_n^-)} \left\{ m u_n^{(1)} \left(\pi_{1,m} G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(T)}\right) + \sum_{p=2}^{m} \frac{m!}{(m-p)!} u_n^{(p)} \left(\pi_{p,m} G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(T)}\right) \right\}.$$

We shall first investigate the linear term $m\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)} u_n^{(1)} \left(\pi_{1,m} G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(T)}\right)$. Notice that

$$m\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)} u_n^{(1)} \left(\pi_{1,m} G_{\varphi,\mathbf{t}}^{(T)}\right) = \frac{m\sqrt{\tilde{\phi}(\mathbf{D}_n^-)}}{\sqrt{n}} \sum_{i=1}^{n} \pi_{1,m} G_{\varphi,\mathbf{t}}^{(T)}(\mathbf{X}_i, \mathbf{Y}_i)$$

We can write

$$\pi_{1,m} G^{(T)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}(x,y) = \mathbb{E}\left[G^{(T)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}(x,X_2,\ldots,X_m),(y,X_2,\ldots,X_m)\right] - \mathbb{E}\left[G^{(T)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}(\mathbf{X},\mathbf{Y})\right]$$
$$= \mathbb{E}\left[G^{(T)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}(\mathbf{X},\mathbf{Y})\Big|(X_1,Y_1)=(x,y)\right] - \mathbb{E}\left[G^{(T)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}(\mathbf{X},\mathbf{Y})\right].$$

We need to introduce a new function

$$S_{\varphi,\mathbf{t},\mathbf{h}} : \mathcal{X}\times\mathcal{Y} \longrightarrow \mathbb{R}$$
$$(x,y) \longrightarrow m\mathbb{E}\left[\varphi(y)\widetilde{K}\left(\frac{d(\mathbf{t},\mathbf{x})}{\mathbf{h}(\mathbf{t})}\right)\Big|(X_1,Y_1)=(x,y)\right].$$

Hence,

$$m\pi_{1,m}G^{(T)}_{\varphi,\mathbf{t}}(x,y) = \tilde{\phi}^{-1}(\mathbf{D}_n^-)\left(S_{\varphi,\mathbf{t},\mathbf{D}_n^-}(x,y) - \mathbb{E}\left[S_{\varphi,\mathbf{t},\mathbf{D}_n^-}(x,y)\right]\right).$$

The linear term of the process is given by

$$m\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}u_n^{(1)}\left(\pi_{1,m}G^{(T)}_{\varphi,\mathbf{t}}\right) = \frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{i=1}^{n}\left\{S_{\varphi,\mathbf{t},\mathbf{D}_n^-}(X_i,Y_i) - \mathbb{E}\left[S_{\varphi,\mathbf{t},\mathbf{D}_n^-}(X_i,Y_i)\right]\right\}$$
$$:= \alpha_n(S_{\varphi,\mathbf{t},\mathbf{D}_n^-}).$$

Therefore, the linear term of the U-process $\{\mu_n(\varphi,\mathbf{t},\mathbf{D}_n^-)\}_{\mathscr{F}_m\mathscr{K}^m}$ is an empirical process indexed by the class of functions $\mathcal{S}$ defined by

$$\mathcal{S} = \left\{S_{\varphi,\mathbf{t},\mathbf{h}}(\cdot,\cdot)\ \varphi\in\mathscr{F}_m, \mathbf{t}=(t_1,\ldots,t_m)\in\mathcal{X}^m\right\},$$

therefore, its weak convergence may be established in a similar way as in the proof of Theorem 4.5. It's clear that $\mathcal{S}\subset m\mathcal{G}^{(1)}$. We consider now the nonlinear part, we have to show that

$$\left\|\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}u_n^{(p)}\left(\pi_{k,m}G^{(T)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\right)\right\|_{\mathscr{F}_m\mathscr{K}^m} \xrightarrow{\mathbb{P}} 0, \quad \text{for} \quad 2\le p\le m.$$

This is a consequence of the Lemma 8.16. Note that the choice of the number and size of the blocks must be made in such a way that the terms I – VI converge to 0. We need to prove that

$$\mathbb{P}\left\{\left\|\mu_n^{(R)}(\varphi,\mathbf{t},\mathbf{D}_n^-)\right\|_{\mathscr{F}_m\mathscr{K}^m} > \lambda\right\} \to 0 \ \text{ as } \ n\to\infty.$$

Again, for clarity purposes, we restrict ourselves to $m=2$. We have

$$\begin{aligned}
&\mu_n^{(R)}(\varphi,\mathbf{t},\mathbf{D}_n^-)\\
&= \sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}\left\{u_n^{(R)}(\varphi,\mathbf{t},\mathbf{D}_n^-) - \mathbb{E}\left(u_n^{(R)}(\varphi,\mathbf{t},\mathbf{D}_n^-)\right)\right\}\\
&= \frac{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}{n(n-1)}\sum_{i\neq j}^{n}\left\{G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((X_i,X_j),(Y_i,Y_j)\right) - \mathbb{E}\left[G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((X_i,X_j),(Y_i,Y_j)\right)\right]\right\}\\
&\leqslant \frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((X_i,X_j),(Y_i,Y_j)\right) - \mathbb{E}\left[G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((X_i,X_j),(Y_i,Y_j)\right)\right]\right\}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)\right]\right\}\\
&+2\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p=1}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)\right]\right\}\\
&+2\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p=1}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\leqslant 1}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)\right]\right\}\\
&+\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in T_p^{(U)}}\sum_{j\in T_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)\right]\right\}\\
&+\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in T_p^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)\right]\right\}\\
=:&\quad \mathrm{I}'+\mathrm{II}'+\mathrm{III}'+\mathrm{IV}'+\mathrm{V}'+\mathrm{VI}'.
\end{aligned}
$$

We will use blocking arguments and treat the resulting terms. We start by considering the first I′. We have

$$
\begin{aligned}
&\mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(X_i,X_j\right),\left(Y_i,Y_j\right)\right)\right]\right\}\right\|_{\mathscr{F}_2\mathscr{K}^2}>\delta\right\}\\
\leqslant\;&\mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(\varsigma_i,\varsigma_j\right),\left(\zeta_i,\zeta_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(\varsigma_i,\varsigma_j\right),\left(\zeta_i,\zeta_j\right)\right)\right]\right\}\right\|_{\mathscr{F}_2\mathscr{K}^2}>\delta\right\}\\
&+2\upsilon_n\beta_{b_n}.
\end{aligned}
$$

Notice that (4.13) readily implies that $\upsilon_n\beta_{b_n}\to 0$ and recall that for all $\varphi\in\mathscr{F}_m$

$$
\mathbf{x},\mathbf{t}\in\mathcal{X}^2,\mathbf{y}\in\mathcal{Y}^2:\kappa_2^2\mathbb{1}_{\{d(\mathbf{x},\mathbf{t})\leqslant h_K\}}F(\mathbf{y})\geqslant\varphi(\mathbf{y})\widetilde{K}\left(\frac{d\left(\mathbf{x},\mathbf{t}\right)}{h_K}\right).
$$

By the symmetry of the function $F(\cdot)$, it holds that

$$
\begin{aligned}
&\left\|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(\varsigma_i,\varsigma_j\right),\left(\zeta_i,\zeta_j\right)\right)-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left(\left(\varsigma_i,\varsigma_j\right),\left(\zeta_i,\zeta_j\right)\right)\right]\right\}\right\|_{\mathscr{F}_2\mathscr{K}^2}\\
&\quad\leqslant\left|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\left\{\kappa^2F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right.\right.\\
&\qquad\qquad\left.\left.-\mathbb{E}\left[\kappa^2F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right]\right\}\right|.
\end{aligned}\tag{8.94}
$$

We use while maintaining order, Chebyshev's inequality and Hoeffding's trick, then

$$\mathbb{P}\left\{\left|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\left\{\kappa^2F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right.\right.\right.$$

$$\left.\left.\left.-\mathbb{E}\left[\kappa^2F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right]\right\}\right|>\delta\right\}$$

$$\leqslant \delta^{-2}n^{-1}\tilde{\phi}^{-1}(\mathbf{D}_n^-)Var\left(\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\kappa^2F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right)$$

$$\leqslant c_2\upsilon_n\delta^{-2}n^{-1}\tilde{\phi}^{-1}(\mathbf{D}_n^-)Var\left(\sum_{p=1}^{\upsilon_n}\sum_{i,j\in H_p^{(U)}}\kappa^2F(\zeta_i,\zeta_j')\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right)$$

$$\leqslant 2c_2\upsilon_n\delta^{-2}n^{-2}\tilde{\phi}^{-1}(\mathbf{D}_n^-)\mathbb{E}\left[\left(\kappa^2F(\zeta_1,\zeta_2)\right)^2\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right]. \tag{8.95}$$

Under **(C.6.)**, we have for each $\lambda > 0$

$$c_2\upsilon_n\delta^{-2}n^{-2}\tilde{\phi}^{-1}(\mathbf{D}_n^-)\mathbb{E}\left[\left(\kappa^2F(\zeta_1,\zeta_2)\right)^2\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right]$$

$$= c_2\upsilon_n\delta^{-2}n^{-2}\tilde{\phi}^{-1}(\mathbf{D}_n^-)\int_0^\infty\mathbb{P}\left\{\left(\kappa^2F(\zeta_1,\zeta_2)\right)^2\mathbb{1}_{\left\{\kappa^2F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\geqslant t\right\}dt$$

$$= c_2\upsilon_n\delta^{-2}n^{-2}\tilde{\phi}^{-1}(\mathbf{D}_n^-)\int_0^{\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}}\mathbb{P}\left\{\kappa^2F>\lambda\left(n\big/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}dt$$

$$+ c_2\upsilon_n\delta^{-2}n^{-2}\tilde{\phi}^{-1}(\mathbf{D}_n^-)\int_{\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}}^\infty\mathbb{P}\left\{\left(\kappa^2F\right)^2>t\right\}dt,$$

which tends to 0 as $n\to\infty$. The terms II′, V′ and VI′ are treated in the same way as the first, except that for II′, VI′ we do not need to apply Hoeffding's trick because our variables $\{\zeta_i,\zeta_j\}_{i,j\in H_p^{(U)}}$ $\left(\text{or } \{\zeta_i,\zeta_j\}_{i,j\in T_p^{(U)}}\right.$ for VI′$\left.\right)$ are in the same blocks, and for the term IV′ we deduce its study from those of I′ and III′. Let us consider the term III′. As for the truncated part, we have

$$\mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p=1}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left((X_i,X_j),(Y_i,Y_j)\right)\right.\right.\right.$$

$$\left.\left.\left.-\mathbb{E}\left[G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left((X_i,X_j),(Y_i,Y_j)\right)\right]\right\}\right\|_{\mathscr{F}_2\mathscr{K}^2}>\delta\right\}$$

$$\leqslant \mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{p=1}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\tilde{\phi}(\mathbf{D}_n^-)\left\{G_{\varphi,\mathbf{t},\mathbf{D}_n^-}^{(R)}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right.\right.\right.$$

$$-\mathbb{E}\left[G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right]\Big\}\Bigg\|_{\mathscr{F}_2\mathscr{K}^2} > \delta\Bigg\}$$

$$+\frac{\upsilon_n^2 a_n b_n \beta_{a_n}}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}. \tag{8.96}$$

We also have

$$\mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\left\{G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)-\mathbb{E}\left[G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right]\right\}\right\|_{\mathscr{F}_2\mathscr{K}^2} > \delta\right\}$$

$$\leqslant \mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\left\{G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)-\mathbb{E}\left[G^{(R)}_{\varphi,\mathbf{t},\mathbf{D}_n^-}\left((\varsigma_i,\varsigma_j),(\zeta_i,\zeta_j)\right)\right]\right\}\right\|_{\mathscr{F}_2\mathscr{K}^2} > \delta\right\}.$$

Since the Eq (8.94) is still satisfied, the problem is reduced to

$$\mathbb{P}\Bigg\{\Bigg|\frac{1}{\sqrt{n\tilde{\phi}(\mathbf{D}_n^-)}}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\Bigg\{\kappa^2 F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2 F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}$$

$$-\mathbb{E}\left[\kappa^2 F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2 F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right]\Bigg\}\Bigg| > \delta\Bigg\}$$

$$\leqslant \delta^{-2}n^{-1}\tilde{\phi}(\mathbf{D}_n^-)Var\left(\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\kappa^2 F(\zeta_i,\zeta_j)\mathbb{1}_{\left\{\kappa^2 F>\lambda\left(n/\tilde{\phi}(\mathbf{D}_n^-)\right)^{1/2(p-1)}\right\}}\right),$$

we follow the same procedure as in (8.95). The rest has just been shown to be asymptotically negligible, so the process $\{\mu_n(\varphi,\mathbf{t},\mathbf{D}_n^-)\}_{\mathscr{F}_m\mathscr{K}^m}$ converges in law to a Gaussian process which has a version with uniformly bounded and uniformly continuous paths with respect to $\|\cdot\|_2-$norm. By repeating the same steps, this also holds true, for the process $\{\mu_n(\varphi,\mathbf{t},\mathbf{D}_n^+)\}_{\mathscr{F}_m\mathscr{K}^m}$. Consequently, by (8.93) it follows that the process $\{\mu_n(\varphi,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t}))\}_{\mathscr{F}_m\mathscr{K}^m}$ also converges in law to a Gaussian process which has a version with uniformly bounded and uniformly continuous paths with respect to $\|\cdot\|_2-$norm. In a similar way, we treat $\{\mu_n(1,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t}))\}_{\mathscr{K}^m}$, and about $\mathbb{E}(u_n(\varphi,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t})))$, $u_n(1,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t}))$ and $\mathbb{E}(u_n(1,\mathbf{t},\mathbf{h}_{n,k}(\mathbf{t})))$ the treatment is done as in the proof of Theorem 4.5.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors would like to thank the Editor-in-Chief, an Associate-Editor, and four anonymous referees for their constructive remarks, which resulted in a substantial improvement of the work's original form and a more sharply focused presentation.

The second author gratefully acknowledges the funding received towards his PhD from the Algerian government PhD fellowship.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1 J. Abrevaya, W. Jiang, A nonparametric approach to measuring and testing curvature, *J. Bus. Econ. Stat.*, **23** (2005), 1–19. https://doi.org/10.1198/073500104000000316

2 H. Akaike, An approximation to the density function, *Ann. Inst. Stat. Math.*, **6** (1954), 127–132. https://doi.org/10.1007/BF02900741

3 I. M. Almanjahie, S. Bouzebda, Z. C. Elmezouar, A. Laksaci, The functional *k*NN estimator of the conditional expectile: uniform consistency in number of neighbors, *Statist. Risk Model.*, **38** (2022), 47–63. https://doi.org/10.1515/strm-2019-0029

4 I. M. Almanjahie, S. Bouzebda, Z. Kaid, A. Laksaci, Nonparametric estimation of expectile regression in functional dependent data, *J. Nonparametr. Stat.*, **34** (2022), 250–281. https://doi.org/10.1080/10485252.2022.2027412

5 I. M. Almanjahie, S. Bouzebda, Z. Kaid, A. Laksaci, The local linear functional *k*NN estimator of the conditional expectile: Uniform consistency in number of neighbors, *Metrika*, **2024** (2024), 1–24. https://doi.org/10.1007/s00184-023-00942-0

6 G. Aneiros, R. Cao, R. Fraiman, C. Genest, P. Vieu, Recent advances in functional data analysis and high-dimensional statistics, *J. Multivariate Anal.*, **170** (2019), 3–9. https://doi.org/10.1016/j.jmva.2018.11.007

7 A. Araujo, E. Giné, *The central limit theorem for real and Banach valued random variables*, New York: John Wiley & Sons, 1980.

8 M. A. Arcones, B. Yu. Central limit theorems for empirical and U-processes of stationary mixing sequences, *J. Theor. Probab.*, **7** (1994), 47–71. https://doi.org/10.1007/BF02213360

9 M. A. Arcones, A Bernstein-type inequality for U-statistics and U-processes, *Stat. Probabil. Lett.*, **22** (1995), 239–247. https://doi.org/10.1016/0167-7152(94)00072-G

10 M. A. Arcones, E. Giné, Limit theorems for U-processes, *Ann. Probab.*, **21** (1993), 1494–1542. https://doi.org/10.1214/aop/1176989128

11 M. A. Arcones, Y. Wang, Some new tests for normality based on U-processes, *Stat. Probabil. Lett.*, **76** (2006), 69–82. https://doi.org/10.1016/j.spl.2005.07.003

12 M. Attouch, A. Laksaci, F. Rafaa, On the local linear estimate for functional regression: uniform in bandwidth consistency, *Commun. Stat. Theor. M.*, **48** (2019), 1836–1853. https://doi.org/10.1080/03610926.2018.1440308

13 A. K. Basu, A. Kundu, Limit distribution for conditional U-statistics for dependent processes, *Calcutta Statistical Association Bulletin*, **52** (2002), 381–407. https://doi.org/10.1177/0008068320020522

14 D. Z. Bello, M. Valk, G. B. Cybis, Towards U-statistics clustering inference for multiple groups, *J. Stat. Comput. Sim.*, **94** (2024), 204–222. https://doi.org/10.1080/00949655.2023.2239978 .

15 N. Berrahou, S. Bouzebda, L. Douge, Functional uniform-in-bandwidth moderate deviation principle for the local empirical processes involving functional data, *Math. Methods Statist.*, **33** (2024), 1–43.

16 P. K. Bhattachary, Y. P. Mack, Weak convergence of k-NN density and regression estimators with varying k and applications, *Ann. Statist.*, **15** (1987), 976–994. https://doi.org/10.1214/aos/1176350487

17 G. Biau, L. Devroye, *Lectures on the nearest neighbor method*, Cham: Springer, 2015. https://doi.org/10.1007/978-3-319-25388-6

18 V. I. Bogachev, *Gaussian measures (Mathematical surveys and monographs)*, Providence: American Mathematical Society, 1998.

19 E. Bolthausen, The Berry-Esseen theorem for functionals of discrete Markov chains, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **54** (1980), 59–73. https://doi.org/10.1007/BF00535354

20 Y. V. Borovskikh, *U-Statistics in Banach spaces*, Boston: De Gruyter, 1996. https://doi.org/10.1515/9783112313954

21 D. Bosq, *Linear processes in function spaces*, New York: Springer-Verlag, 2000. https://doi.org/10.1007/978-1-4612-1154-9

22 B. Feriel, O. S. Elias, Nonparametric local linear estimation of the relative error regression function for twice censored data, *Stat. Probabil. Lett.*, **178** (2021), 109185. https://doi.org/10.1016/j.spl.2021.109185

23 F. Bouhadjerad, E. O. Saïd, Strong consistency of the local linear relative regression estimator for censored data, *Opuscula Math.*, **42** (2022), 805–832. https://doi.org/10.7494/OpMath.2022.42.6.805

24 F. Bouhadjera, M. Lemdani, E, O. Saïd, Strong uniform consistency of the local linear relative error regression estimator under left truncation, *Stat. Papers*, **64** (2023), 421–447. https://doi.org/10.1007/s00362-022-01325-9

25 S. Bouzebda, On the strong approximation of bootstrapped empirical copula processes with applications, *Math. Meth. Stat.*, **21** (2012), 153–188. https://doi.org/10.3103/S1066530712030015

26 S. Bouzebda, Asymptotic properties of pseudo maximum likelihood estimators and test in semi-parametric copula models with multiple change points, *Math. Meth. Stat.*, **23** (2014), 38–65. https://doi.org/10.3103/S1066530714010037

27 S. Bouzebda, B. Nemouchi, Central limit theorems for conditional empirical and conditional U-processes of stationary mixing sequences, *Math. Meth. Stat.*, **28** (2019), 169–207. https://doi.org/10.3103/S1066530719030013

28 S. Bouzebda, N. Taachouche, Rates of the strong uniform consistency with rates for conditional U-statistics estimators with general kernels on manifolds, *Math. Meth. Stat.*, **33** (2024), 1–55.

29 S. Bouzebda, T. Zari, Strong approximation of multidimensional $\mathbb{P}$-$\mathbb{P}$ plots processes by Gaussian processes with applications to statistical tests, *Math. Meth. Stat.*, **23** (2014), 210–238. https://doi.org/10.3103/S1066530714030041

30 S. Bouzebda, M. Chaouch, N. Laïb, Limiting law results for a class of conditional mode estimates for functional stationary ergodic data, *Math. Meth. Stat.*, **25** (2016), 168–195. https://doi.org/10.3103/S1066530716030029 .

31 S. Bouzebda, Strong approximation of the smoothed Q-Q processes, *Far East Journal of Theoretical Statistics*, **31** (2010), 169–191.

32 S. Bouzebda, General tests of independence based on empirical processes indexed by functions, *Stat. Methodol.*, **21** (2014), 59–87. https://doi.org/10.1016/j.stamet.2014.03.001

33 S. Bouzebda, On the weak convergence and the uniform-in-bandwidth consistency of the general conditional U-processes based on the copula representation: multivariate setting, *Hacet. J. Math. Stat.*, **52** (2023), 1303–1348. https://doi.org/10.15672/hujms.1134334

34 S. Bouzebda, General tests of conditional independence based on empirical processes indexed by functions, *Jpn. J. Stat. Data Sci.*, **6** (2023), 115–177. https://doi.org/10.1007/s42081-023-00193-3

35 S. Bouzebda, M. Chaouch, Uniform limit theorems for a class of conditional Z-estimators when covariates are functions, *J. Multivariate Anal.*, **189** (2022), 104872. https://doi.org/10.1016/j.jmva.2021.104872

36 K. Chokri, S. Bouzebda, Uniform-in-bandwidth consistency results in the partially linear additive model components estimation, *Commun. Stat. Theor. M.*, **2023** (2023), 2153605. https://doi.org/10.1080/03610926.2022.2153605

37 S. Bouzebda, I. Elhattab, A strong consistency of a nonparametric estimate of entropy under random censorship, *CR Math.*, **347** (2009), 821–826. https://doi.org/10.1016/j.crma.2009.04.021

38 S. Bouzebda, I. Elhattab, Uniform-in-bandwidth consistency for kernel-type estimators of Shannon's entropy, *Electron. J. Stat.*, **5** (2011), 440–459. https://doi.org/10.1214/11-EJS614

39 S. Bouzebda, A. A. Ferfache, Asymptotic properties of M-estimators based on estimating equations and censored data in semi-parametric models with multiple change points, *J. Math. Anal. Appl.*, **497** (2021), 124883. https://doi.org/10.1016/j.jmaa.2020.124883

40 S. Bouzebda, A. A. Ferfache, Functional central limit theorems for triangular arrays of function-indexed U-processes under uniformly integrable entropy conditions, submitted for publication.

41 S. Bouzebda, A. A. Ferfache, Asymptotic properties of semiparametric M-estimators with multiple change points, *Phys. A*, **609** (2023), 128363. https://doi.org/10.1016/j.physa.2022.128363

42 S. Bouzebda, 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), 452–509. https://doi.org/10.1080/10485252.2020.1759597

43 S. Bouzebda, B. Nemouchi, Weak-convergence of empirical conditional processes and conditional U-processes involving functional mixing data, *Stat. Inference Stoch. Process.*, **26** (2023), 33–88. https://doi.org/10.1007/s11203-022-09276-6

44 S. Bouzebda, 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** (2022), 431–533. https://doi.org/10.1007/s42081-022-00161-3

45 S. Bouzebda, A. Nezzal, Asymptotic properties of conditional U-statistics using delta sequences, *Commun. Stat. Theor. M.*, **2023** (2023), 2179887. https://doi.org/10.1080/03610926.2023.2179887

46 S. Bouzebda, I. Soukarieh, Renewal type bootstrap for U-process Markov chains, *Markov Process. Relat.*, **28** (2022), 673–735.

47 S. Bouzebda, I. Soukarieh, Non-parametric conditional U-processes for locally stationary functional random fields under stochastic sampling design, *Mathematics*, **11** (2023), 16. https://doi.org/10.3390/math11010016 .

48 S. Bouzebda, I. Soukarieh, Limit theorems for a class of processes generalizing the U-empirical process, in press.

49 S. Bouzebda, N. Taachouche, On the variable bandwidth kernel estimation of conditional U-statistics at optimal rates in sup-norm, *Phys. A*, **625** (2023), 129000. https://doi.org/10.1016/j.physa.2023.129000 .

50 S. Bouzebda, N. Taachouche, Rates of the strong uniform consistency for the kernel-type regression function estimators with general kernels on manifolds, *Math. Meth. Stat.*, **32** (2023), 27–80. https://doi.org/10.3103/s1066530723010027 .

51 S. Bouzebda, I. Elhattab, C. T. Seck, Uniform in bandwidth consistency of nonparametric regression based on copula representation, *Stat. Probabil. Lett.*, **137** (2018), 173–182. https://doi.org/10.1016/j.spl.2018.01.021

52 S. Bouzebda, I. Elhattab, B. Nemouchi, On the uniform-in-bandwidth consistency of the general conditional U-statistics based on the copula representation, *J. Nonparametr. Stat.*, **33** (2021), 321–358. https://doi.org/10.1080/10485252.2021.1937621

53 S. Bouzebda, A. Laksaci, M. Mohammedi, Single index regression model for functional quasi-associated time series data, *REVSTAT-Stat. J.*, **20** (2022), 605–631. https://doi.org/10.57805/revstat.v20i5.391

54 S. Bouzebda, T. El-hadjali, A. A. Ferfache, Uniform in bandwidth consistency of conditional U-statistics adaptive to intrinsic dimension in presence of censored data, *Sankhya A*, **85** (2023), 1548–1606. https://doi.org/10.1007/s13171-022-00301-7

55 S. Bouzebda, A. Laksaci, M. Mohammedi, The k-nearest neighbors method in single index regression model for functional quasi-associated time series data, *Rev. Mat. Complut.*, **36** (2023), 361–391. https://doi.org/10.1007/s13163-022-00436-z

56 S. Bouzebda, A. Nezzal, T. Zari, Uniform consistency for functional conditional U-statistics using delta-sequences, *Mathematics*, **11** (2023), 161. https://doi.org/10.3390/math11010161

57 J. Bretagnolle, Lois limites du bootstrap de certaines fonctionnelles, *Ann. I. H. Poincare-PR.*, **19** (1983), 281–296.

58 F. Burba, F. Ferraty, P. Vieu, *k*-Nearest neighbour method in functional nonparametric regression, *J. Nonparametr. Stat.*, **21** (2009), 453–469. https://doi.org/10.1080/10485250802668909

59 L. Chen, A. T. K. Wan, S. Zhang, Y. Zhou, Distributed algorithms for U-statistics-based empirical risk minimization, *J. Mach. Learn. Res.*, **24** (2023), 1–43.

60 Z. Chikr-Elmezouar, I. M. Almanjahie, A. Laksaci, M. Rachdi, FDA: strong consistency of the *k*NN local linear estimation of the functional conditional density and mode, *J. Nonparametr. Stat.*, **31** (2019), 175–195. https://doi.org/10.1080/10485252.2018.1538450

61 J. A. Clarkson, C. R. Adams, On definitions of bounded variation for functions of two variables, *T. Am. Math. Soc.*, **35** (1933), 824–854. https://doi.org/10.2307/1989593

62 S. Clémençon, G. Lugosi, N. Vayatis, Ranking and empirical minimization of U-statistics, *Ann. Statist.*, **36** (2008), 844–874. https://doi.org/10.1214/009052607000000910

63 S. Clémençon, I. Colin, A. Bellet, Scaling-up empirical risk minimization: optimization of incomplete U-statistics, *J. Mach. Learn. Res.*, **17** (2016), 76.

64 G. Collomb, Estimation de la régression par la méthode des *k* points les plus proches avec noyau: quelques propriétés de convergence ponctuelle, In: *Statistique non paramétrique asymptotique*, Berlin: Springer, 1980, 159–175. https://doi.org/10.1007/BFb0097428

65 G. B. Cybis, M. Valk, S. R. C. Lopes, Clustering and classification problems in genetics through U-statistics, *J. Stat. Comput. Sim.*, **88** (2018), 1882–1902. https://doi.org/10.1080/00949655.2017.1374387

66 Y. A. Davydov, Mixing conditions for Markov chains, *Theor. Probab. Appl.*, **18** (1974), 321–338. https://doi.org/10.1137/1118033

67 V. H. de la Peña, Decoupling and Khintchine's inequalities for U-statistics, *Ann. Probab.*, **20** (1992), 1877–1892. https://doi.org/10.1214/aop/1176989533

68 V. H. de la Peña, E. Giné, *Decoupling*, New York: Springer, 1999. https://doi.org/10.1007/978-1-4612-0537-1

69 J. Dedecker, S. Louhichi, Maximal inequalities and empirical central limit theorems, In: *Empirical process techniques for dependent data*, Boston: Birkhäuser, 2002, 137–159. https://doi.org/10.1007/978-1-4612-0099-4_3

70 P. Deheuvels, One bootstrap suffices to generate sharp uniform bounds in functional estimation, *Kybernetika (Prague)*, **47** (2011), 855–865.

71 P. Deheuvels, D. M. Mason, General asymptotic confidence bands based on kernel-type function estimators, *Stat. Infer. Stoch. Pro.*, **7** (2004), 225–277. https://doi.org/10.1023/B:SISP.0000049092.55534.af

72 H. Dehling, M. Wendler, Central limit theorem and the bootstrap for U-statistics of strongly mixing data, *J. Multivariate Anal.*, **101** (2010), 126–137. https://doi.org/10.1016/j.jmva.2009.06.002

73 K. Dehnad, Density estimation for statistics and data analysis, *Technometrics*, **29** (1987), 495–495. https://doi.org/10.1080/00401706.1987.10488295

74 J. Demongeot, A. Hamie, A. Laksaci, M. Rachdi, Relative-error prediction in nonparametric functional statistics: theory and practice, *J. Multivariate Anal.*, **146** (2016), 261–268. https://doi.org/10.1016/j.jmva.2015.09.019

75 L. Devroye, *A course in density estimation*, Boston: Birkhäuser Boston Inc., 1987.

76 L. Devroye, G. Lugosi, *Combinatorial methods in density estimation*, New York: Springer-Verlag, 2001. https://doi.org/10.1007/978-1-4613-0125-7

77 L. Devroye, L. Györfi, A. Krzyzak, G. Lugosi, On the strong universal consistency of nearest neighbor regression function estimates, *Ann. Statist.*, **22** (1994), 1371–1385. https://doi.org/10.1214/aos/1176325633

78 J. Dony, U. Einmahl, Uniform in bandwidth consistency of kernel regression estimators at a fixed point, *Inst. Math. Stat. (IMS) Collect.*, **2009** (2009), 308–325. https://doi.org/10.1214/09-IMSCOLL520

79 J. Dony, D. M. Mason, Uniform in bandwidth consistency of conditional U-statistics, *Bernoulli*, **14** (2008), 1108–1133. https://doi.org/10.3150/08-BEJ136

80 R. M. Dudley, The sizes of compact subsets of Hilbert space and continuity of Gaussian processes, *J. Funct. Anal.*, **1** (1967), 290–330. https://doi.org/10.1016/0022-1236(67)90017-1

81 R. M. Dudley, A course on empirical processes, In: *École d'été de probabilités de Saint-Flour, XII-1982*, Berlin: Springer, 1984, 1–142. https://doi.org/10.1007/BFb0099432

82 R. M. Dudley, *Uniform central limit theorems*, Cambridge: Cambridge University Press, 1999. https://doi.org/10.1017/CBO9780511665622

83 E. B. Dynkin, A. Mandelbaum, Symmetric statistics, poisson point processes, and multiple wiener integrals, *Ann. Statist.*, **11** (1983), 739–745. https://doi.org/10.1214/aos/1176346241

84 E. Eberlein, Weak convergence of partial sums of absolutely regular sequences, *Stat. Probabil. Lett.*, **2** (1984), 291–293. https://doi.org/10.1016/0167-7152(84)90067-1

85 S. Efromovich, *Nonparametric curve estimation*, New York: Springer, 1999. https://doi.org/10.1007/b97679

86 P. P. B. Eggermont, V. N. LaRiccia, *Maximum penalized likelihood estimation*, New York: Springer, 2001. https://doi.org/10.1007/978-1-0716-1244-6

87 U. Einmahl, D. M. Mason, An empirical process approach to the uniform consistency of kernel-type function estimators, *J. Theor. Probab.*, **13** (2000), 1–37. https://doi.org/10.1023/A:1007769924157 .

88 U. Einmahl, D. M. Mason, Uniform in bandwidth consistency of kernel-type function estimators, *Ann. Statist.*, **33** (2005), 1380–1403. https://doi.org/10.1214/009053605000000129

89 M. Ezzahrioui, E. Ould-Saïd, Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data, *J. Nonparametr. Stat.*, **20** (2008), 3–18. https://doi.org/10.1080/10485250701541454

90 L. Faivishevsky, J. Goldberger, ICA based on a smooth estimation of the differential entropy, In: *Proceedings of the 21st international conference on neural information processing systems*, New York: Curran Associates, Inc., 2008, 433–440.

91 F. Ferraty, P. Vieu, *Nonparametric functional data analysis*, New York: Springer, 2006. https://doi.org/10.1007/0-387-36620-2

92 F. Ferraty, A. Mas, P. Vieu, Nonparametric regression on functional data: inference and practical aspects, *Aust. N. Z. J. Stat.*, **49** (2007), 267–286. https://doi.org/10.1111/j.1467-842X.2007.00480.x

93 F. Ferraty, A. Laksaci, A. Tadj, P. Vieu, Rate of uniform consistency for nonparametric estimates with functional variables, *J. Stat. Plan. Infer.*, **140** (2010), 335–352. https://doi.org/10.1016/j.jspi.2009.07.019

94 A. A. Filippova, Mises theorem on the limit behaviour of functionals derived from empirical distribution functions, *Dokl. Akad. Nauk SSSR*, **129** (1959), 44–47. https://doi.org/10.1137/1107003

95 E. Fix, J. L. J. Hodges, Discriminatory analysis-nonparametric discrimination: consistency properties, *Technical Report Project 21-49-004, Report 4*, USAF School of Aviation Medicine, Randolph Field, Texas, 1951.

96 E. Fix, J. L. J. Hodges, Discriminatory analysis–nonparametric discrimination: consistency properties, *Int. Stat. Rev.*, **57** (1989), 238–247. https://doi.org/10.2307/1403797

97 E. W. Frees, Infinite order U-statistics, *Scand. J. Stat.*, **16** (1989), 29–45.

98 K.-A. Fu, An application of U-statistics to nonparametric functional data analysis, *Commun. Stat. Theor. M.*, **41** (2012), 1532–1542. https://doi.org/10.1080/03610926.2010.526747

99 T. Gasser, P. Hall, B. Presnell, Nonparametric estimation of the mode of a distribution of random curves, *J. R. Stat. Soc. B*, **60** (1998), 681–691. https://doi.org/10.1111/1467-9868.00148

100 S. Ghosal, A. Sen, A. W. van der Vaart, Testing monotonicity of regression, *Ann. Statist.*, **28** (2000), 1054–1082. https://doi.org/10.1214/aos/1015956707

101 E. Giné, D. M. Mason, Laws of the iterated logarithm for the local U-statistic process, *J. Theor. Probab.*, **20** (2007), 457–485. https://doi.org/10.1007/s10959-007-0067-0

102 E. Giné, J. Zinn, Some limit theorems for empirical processes, *Ann. Probab.*, **12** (1984), 929–989. https://doi.org/10.1214/aop/1176993138

103 H. L. Gray, N.-F. Zhang, W. A. Woodward, On generalized fractional processes, *J. Time Ser. Anal.*, **10** (1989), 233–257. https://doi.org/10.1111/j.1467-9892.1989.tb00026.x

104 L. Györfi, The rate of convergence of k-nn regression estimation and classification, *IEEE T. Inform. Theory*, **27** (1981), 362–364. https://doi.org/10.1109/TIT.1981.1056344

105 P. Hall, Asymptotic properties of integrated square error and cross-validation for kernel estimation of a regression function, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **67** (1984), 175–196. https://doi.org/10.1007/BF00535267

106 P. R. Halmos, The theory of unbiased estimation, *Ann. Math. Statist.*, **17** (1946), 34–43. https://doi.org/10.1214/aoms/1177731020

107 F. Han, An exponential inequality for U-statistics under mixing conditions, *J. Theor. Probab.*, **31** (2018), 556–578. https://doi.org/10.1007/s10959-016-0722-4

108 W. Härdle, *Applied nonparametric regression*, Cambridge: Cambridge University Press, 1990. https://doi.org/10.1017/CCOL0521382483

109 W. Härdle, J. S. Marron, Optimal bandwidth selection in nonparametric regression function estimation, *Ann. Statist.*, **13** (1985), 1465–1481. https://doi.org/10.1214/aos/1176349748

110 G. H. Hardy, On double fourier series and especially those which represent the double zeta-function with real and incommensurable parameters, *Quart. J. Math*, **37** (1905), 53–79.

111 M. Harel, M. L. Puri, Conditional U-statistics for dependent random variables, *J. Multivariate Anal.*, **57** (1996), 84–100. https://doi.org/10.1006/jmva.1996.0023

112 C. Heilig, D. Nolan, Limit theorems for the infinite-degree U-process, *Stat. Sinica*, **11** (2001), 289–302.

113 L. Heinrich, Bounds for the absolute regularity coefficient of a stationary renewal process, *Yokohama Math. J.*, **40** (1992), 25–33.

114 E. W. Hobson, *The theory of functions of a real variable and the theory of Fourier's series. Vol. II*, New York: Dover Publications, Inc., 1958.

115 W. Hoeffding, A class of statistics with asymptotically normal distribution, *Ann. Math. Statist.*, **19** (1948), 293–325. https://doi.org/10.1214/aoms/1177730196

116 L. Horváth, P. Kokoszka, *Inference for functional data with applications*, New York: Springer, 2012. https://doi.org/10.1007/978-1-4614-3655-3

117 P. J. Huber, Robust estimation of a location parameter, *Ann. Math. Statist.*, **35** (1964), 73–101. https://doi.org/10.1214/aoms/1177703732

118 I. A. Ibragimov, V. N. Solev, A condition for regularity of a Gaussian stationary process, *Soviet Math. Dokl.*, **10** (1969), 371–375.

119 S. Jadhav, S. Ma, An association test for functional data based on Kendall's Tau, *J. Multivariate Anal.*, **184** (2021), 104740. https://doi.org/10.1016/j.jmva.2021.104740

120 S. Janson, A functional limit theorem for random graphs with applications to subgraph count statistics, *Random Struct. Algor.*, **1** (1990), 15–37. https://doi.org/10.1002/rsa.3240010103

121 S. Janson, Asymptotic normality for m-dependent and constrained U-statistics, with applications to pattern matching in random strings and permutations, *Adv. Appl. Probab.*, **55** (2023), 841–894. https://doi.org/10.1017/apr.2022.51

122 E. Joly, G. Lugosi, Robust estimation of U-statistics, *Stoch. Proc. Appl.*, **126** (2016), 3760–3773. https://doi.org/10.1016/j.spa.2016.04.021

123 M. C. Jones, H. Park, K. Shin, S. K. Vines, S. Jeong, Relative error prediction via kernel regression smoothers, *J. Stat. Plan. Infer.*, **138** (2008), 2887–2898. https://doi.org/10.1016/j.jspi.2007.11.001

124 L. Kara, A. Laksaci, M. Rachdi, P. Vieu, Data-driven kNN estimation in nonparametric functional data analysis, *J. Multivariate Anal.*, **153** (2017), 176–188. https://doi.org/10.1016/j.jmva.2016.09.016

125 L. Kara-Zaitri, A. Laksaci, M. Rachdi, P. Vieu, Uniform in bandwidth consistency for various kernel estimators involving functional data, *J. Nonparametr. Stat.*, **29** (2017), 85–107. https://doi.org/10.1080/10485252.2016.1254780

126 H. A. Karlsen, D. Tjøstheim, Nonparametric estimation in null recurrent time series, *Ann. Statist.*, **29** (2001), 372–416. https://doi.org/10.1214/aos/1009210546

127 M. G. Kendall, A new measure of rank correlation, *Biometrika*, **30** (1938), 81–93. https://doi.org/10.2307/2332226

128 I. Kim, A. Ramdas, Dimension-agnostic inference using cross U-statistics, *Bernoulli*, **30** (2024), 683–711. https://doi.org/10.3150/23-bej1613

129 R. Koenker, G. Bassett, Regression quantiles, *Econometrica*, **46** (1978), 33–50. https://doi.org/10.2307/1913643

130 P. Kokoszka, M. Reimherr, *Introduction to functional data analysis*, Boca Raton: Chapman and Hall/CRC Press, 2017. https://doi.org/10.1201/9781315117416

131 A. N. Kolmogorov, V. M. Tikhomirov, ε-entropy and ε-capacity of sets in function spaces, *Uspekhi Mat. Nauk*, **14** (1959), 3–86.

132 M. R. Kosorok, *Introduction to empirical processes and semiparametric inference*, New York: Springer, 2008. https://doi.org/10.1007/978-0-387-74978-5

133 M. Krause, Über mittelwertsätze im Gebiete der doppelsummen und doppelintegrale, *Leipz. Ber.*, **55** (1903), 239–263.

134 N. L. Kudraszow, P. Vieu, Uniform consistency of *k*NN regressors for functional variables, *Stat. Probabil. Lett.*, **83** (2013), 1863–1870. https://doi.org/10.1016/j.spl.2013.04.017

135 T. Laloë, A *k*-nearest neighbor approach for functional regression, *Stat. Probabil. Lett.*, **78** (2008), 1189–1193. https://doi.org/10.1016/j.spl.2007.11.014

136 T. L. Minh, *U*-statistics on bipartite exchangeable networks, *ESAIM Probab. Stat.*, **27** (2023), 576–620. https://doi.org/10.1051/ps/2023010

137 L. LeCam, A remark on empirical measures, In: *A Festschrift for Erich Lehmann in honor of his sixty-fifth birthday*, Belmont: Wadsworth, 1983, 305–327.

138 A. J. Lee, *U-statistics*, New York: Marcel Dekker, Inc., 1990.

139 W. V. Li, Q.-M. Shao, Gaussian processes: inequalities, small ball probabilities and applications, *Handbook of Statistics*, **19** (2001), 533–597. https://doi.org/10.1016/S0169-7161(01)19019-X

140 F. Lim, V. M. Stojanovic, On U-statistics and compressed sensing I: Non-asymptotic average-case analysis, *IEEE T. Signal Proces.*, **61** (2013), 2473–2485. https://doi.org/10.1109/TSP.2013.2247598

141 N. Ling, S. Meng, P. Vieu, Uniform consistency rate of *k*NN regression estimation for functional time series data, *J. Nonparametr. Stat.*, **31** (2019), 451–468. https://doi.org/10.1080/10485252.2019.1583338

142 N. Ling, G. Aneiros, P. Vieu, *k*NN estimation in functional partial linear modeling, *Stat. Papers*, **61** (2020), 423–444. https://doi.org/10.1007/s00362-017-0946-0

143 Q. Liu, J. Lee, M. Jordan, A kernelized stein discrepancy for goodness-of-fit tests, *The 33rd International Conference on Machine Learning*, New York, USA, 2016, 276–284.

144 D. O. Loftsgaarden, C. P. Quesenberry, A nonparametric estimate of a multivariate density function, *Ann. Math. Statist.*, **36** (1965), 1049–1051. https://doi.org/10.1214/aoms/1177700079

145 Y. P. Mack, Local properties of k-nn regression estimates, *SIAM Journal on Algebraic Discrete Methods*, **2** (1981), 311–323. https://doi.org/10.1137/0602035

146 D. M. Mason, Proving consistency of non-standard kernel estimators, *Stat. Inference Stoch. Process.*, **15** (2012), 151–176. https://doi.org/10.1007/s11203-012-9068-4

147 E. Masry, Nonparametric regression estimation for dependent functional data: asymptotic normality, *Stoch. Proc. Appl.*, **115** (2005), 155–177. https://doi.org/10.1016/j.spa.2004.07.006

148 E. Mayer-Wolf, O. Zeitouni, The probability of small Gaussian ellipsoids and associated conditional moments, *Ann. Probab.*, **21** (1993), 14–24.

149 F. Merlevède, M. Peligrad, E. Rio, A Bernstein type inequality and moderate deviations for weakly dependent sequences, *Probab. Theory Relat. Fields*, **151** (2011), 435–474. https://doi.org/10.1007/s00440-010-0304-9

150 M. Mohammedi, S. Bouzebda, A. Laksaci, On the nonparametric estimation of the functional expectile regression, *CR Math.*, **358** (2020), 267–272. https://doi.org/10.5802/crmath.27

151 M. Mohammedi, S. Bouzebda, A. Laksaci, The consistency and asymptotic normality of the kernel type expectile regression estimator for functional data, *J. Multivariate Anal.*, **181** (2021), 104673. https://doi.org/10.1016/j.jmva.2020.104673

152 M. Mohammedi, S. Bouzebda, A. Laksaci, O. Bouanani, Asymptotic normality of the k-NN single index regression estimator for functional weak dependence data, *Commun. Stat. Theor. M.*, **2022** (2022), 2150823. https://doi.org/10.1080/03610926.2022.2150823

153 E. A. Nadaraja, On estimate regression , *Theor. Probab. Appl.*, **9** (1964), 141–142.

154 E. A. Nadaraya, *Nonparametric estimation of probability densities and regression curves*, Netherlands: Kluwer Academic Publishers, 1989. https://doi.org/10.1007/978-94-009-2583-0

155 W. K. Newey, J. L. Powell, Asymmetric least squares estimation and testing, *Econometrica*, **55** (1987), 819–847. https://doi.org/10.2307/1911031

156 H. Niederreiter, *Random number generation and quasi-Monte Carlo methods*, Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1992. https://doi.org/10.1137/1.9781611970081

157 D. Nolan, D. Pollard, U-processes: rates of convergence, *Ann. Statist.*, **15** (1987), 780–799. https://doi.org/10.1214/aos/1176350374

158 S. Novo, G. Aneiros, P. Vieu, Automatic and location-adaptive estimation in functional single-index regression, *J. Nonparametr. Stat.*, **31** (2019), 364–392. https://doi.org/10.1080/10485252.2019.1567726

159 H. Park, L. A. Stefanski, Relative-error prediction, *Stat. Probabil. Lett.*, **40** (1998), 227–236. https://doi.org/10.1016/S0167-7152(98)00088-1

160 E. Parzen, On estimation of a probability density function and mode, *Ann. Math. Statist.*, **33** (1962), 1065–1076. https://doi.org/10.1214/aoms/1177704472

161 W. Peng, T. Coleman, L. Mentch, Rates of convergence for random forests via generalized U-statistics, *Electron. J. Statist.*, **16** (2022), 232–292. https://doi.org/10.1214/21-ejs1958

162 N. Phandoidaen, S. Richter, Empirical process theory for locally stationary processes, *Bernoulli*, **28** (2022), 453–480. https://doi.org/10.3150/21-bej1351

163 D. Pollard, *Convergence of stochastic processes*, New York: Springer, 1984. https://doi.org/10.1007/978-1-4612-5254-2

164 W. Polonik, Q. Yao, Set-indexed conditional empirical and quantile processes based on dependent data, *J. Multivariate Anal.*, **80** (2002), 234–255. https://doi.org/10.1006/jmva.2001.1988

165 B. L. S. P. Rao, A. Sen, Limit distributions of conditional U-statistics, *J. Theoret. Probab.*, **8** (1995), 261–301. https://doi.org/10.1007/BF02212880

166 M. Rachdi, P. Vieu, Nonparametric regression for functional data: automatic smoothing parameter selection, *J. Stat. Plan. Infer.*, **137** (2007), 2784–2801. https://doi.org/10.1016/j.jspi.2006.10.001

167 J. O. Ramsay, B. W. Silverman. *Applied functional data analysis*, New York: Springer, 2002. https://doi.org/10.1007/b98886

168 J. O. Ramsay, B. W. Silverman, *Functional data analysis*, New York: Springer, 2 Eds., 2005. https://doi.org/10.1007/b98888

169 P. M. Robinson, Large-sample inference for nonparametric regression with dependent errors, *Ann. Statist.*, **25** (1997), 2054–2083. https://doi.org/10.1214/aos/1069362387

170 M. Rosenblatt, A central limit theorem and a strong mixing condition, *P. Nat. Acad. Sci. USA*, **42** (1956), 43–47. https://doi.org/10.1073/pnas.42.1.43

171 M. Rosenblatt, Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.*, **27** (1956), 832–837. https://doi.org/10.1214/aoms/1177728190

172 H. Rubin, R. A. Vitale, Asymptotic distribution of symmetric statistics, *Ann. Statist.*, **8** (1980), 165–170.

173 A. Schick, Y. Wang, W. Wefelmeyer, Tests for normality based on density estimators of convolutions, *Stat. Probabil. Lett.*, **81** (2011), 337–343. https://doi.org/10.1016/j.spl.2010.10.022

174 D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*, New York: John Wiley & Sons Inc., 1992. https://doi.org/10.1002/9780470316849

175 A. Sen, Uniform strong consistency rates for conditional U-statistics, *Sankhyā Ser. A*, **56** (1994), 179–194.

176 R. J. Serfling, *Approximation theorems of mathematical statistics*, New York: John Wiley & Sons, Inc., 1980. https://doi.org/10.1002/9780470316481

177 R. P. Sherman, Maximal inequalities for degenerate U-processes with applications to optimization estimators, *Ann. Statist.*, **22** (1994), 439–459. https://doi.org/10.1214/aos/1176325377

178 Y. Song, X. Chen, K. Kato. Approximating high-dimensional infinite-order U-statistics: statistical and computational guarantees, *Electron. J. Statist.*, **13** (2019), 4794–4848. https://doi.org/10.1214/19-EJS1643

179 I. Soukarieh, S. Bouzebda, Exchangeably weighted bootstraps of general markov U-process, *Mathematics*, **10** (2022), 3745. https://doi.org/10.3390/math10203745

180 I. Soukarieh, S. Bouzebda. Renewal type bootstrap for increasing degree U-process of a Markov chain, *J. Multivariate Anal.*, **195** (2023), 105143. https://doi.org/10.1016/j.jmva.2022.105143

181 I. Soukarieh, S. Bouzebda, Weak convergence of the conditional U-statistics for locally stationary functional time series, *Stat. Inference Stoch. Process.*, **2023** (2023), 1–78. https://doi.org/10.1007/s11203-023-09305-y

182 W. Stute, Conditional U-statistics, *Ann. Probab.*, **19** (1991), 812–825.

183 W. Stute, Almost sure representations of the product-limit estimator for truncated data, *Ann. Statist.*, **21** (1993), 146–156. https://doi.org/10.1214/aos/1176349019

184 W. Stute, L^p-convergence of conditional U-statistics, *J. Multivariate Anal.*, **51** (1994), 71–82. https://doi.org/10.1006/jmva.1994.1050

185 W. Stute, Universally consistent conditional U-statistics, *Ann. Statist.*, **22** (1994), 460–473. https://doi.org/10.1214/aos/1176325378

186 W. Stute, Symmetrized NN-conditional U-statistics, In: *Research developments in probability and statistics*, VSP, Utrecht, 1996, 231–237.

187 J. Su, Z. Yao, C. Li, Y. Zhang. A statistical approach to estimating adsorption-isotherm parameters in gradient-elution preparative liquid chromatography, *Ann. Appl. Stat.*, **17** (2023), 3476–3499. https://doi.org/10.1214/23-aoas1772

188 K. K. Sudheesh, S. Anjana, M. Xie. U-statistics for left truncated and right censored data, *Statistics*, **57** (2023), 900–917. https://doi.org/10.1080/02331888.2023.2217314

189 O. Toussoun, Mémoire sur l'histoire du nil, In: *Mémoires de l'Institut d'Egypte*, Cairo: Institut d'Egypte, 1925.

190 A. W. van der Vaart, *Asymptotic statistics*, Cambridge: Cambridge University Press, 1998. https://doi.org/10.1017/CBO9780511802256 .

191 A. van der Vaart, The statistical work of Lucien Le Cam. *Ann. Statist.*, **30** (2002), 631–682. https://doi.org/10.1214/aos/1028674836

192 A. van der Vaart, H. van Zanten, Bayesian inference with rescaled Gaussian process priors, *Electron. J. Statist.*, **1** (2007), 433–448. https://doi.org/10.1214/07-EJS098

193 A. W. van der Vaart, J. A. Wellner, *Weak convergence and empirical processes*, New York: Springer, 1996. https://doi.org/10.1007/978-1-4757-2545-2

194 V. N. Vapnik, A. Ja. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theor. Probab. Appl.*, **16** (1971), 264–279.

195 G. Vitali, Sui gruppi di punti e sulle funzioni di variabili reali, *Atti Accad. Sci. Torino*, **43** (1908), 75–92.

196 A. G. Vituškin, *O mnogomernyh variaciyah*, Gostehisdat: Moskva, 1955.

197 V. Volkonski, Y. Rozanov, Some limit theorems for random functions. I, *Theor. Probab. Appl.* , **4** (1959), 178–197. https://doi.org/10.1137/1104015

198 R. von Mises, On the asymptotic distribution of differentiable statistical functions, *Ann. Math. Statist.*, **18** (1947), 309–348. https://doi.org/10.1214/aoms/1177730385

199 M. P. Wand, M. C. Jones, *Kernel smoothing*, New York: Chapman and Hall/CRC Press, 1994. https://doi.org/10.1201/b14876

200 G. S. Watson, Smooth regression analysis, *Sankhya: The Indian Journal of Statistics, Series A*, **26** (1964), 359–372.

201 C. Xu, Y. Zhang, Estimating the memory parameter for potentially non-linear and non-Gaussian time series with wavelets, *Inverse Probl.*, **38** (2022), 035004. https://doi.org/10.1088/1361-6420/ac48ca

202 Y. Yajima, On estimation of a regression model with long-memory stationary errors, *Ann. Statist.*, **16** (1988), 791–807. https://doi.org/10.1214/aos/1176350837

203 K. Yoshihara, Limiting behavior of U-statistics for stationary, absolutely regular processes, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **35** (1976), 237–252. https://doi.org/10.1007/BF00532676

204 Y. Zhang, C. Chen, Stochastic asymptotical regularization for linear inverse problems, *Inverse Probl.*, **39** (2023), 015007. https://doi.org/10.1088/1361-6420/aca70f

205 Y. Zhang, Z. Yao, P. Forssén, T. Fornstedt, Estimating the rate constant from biosensor data via an adaptive variational Bayesian approach, *Ann. Appl. Stat.*, **13** (2019), 2011–2042. https://doi.org/10.1214/19-aoas1263

A. Proofs of lemmas

Proof of Lemma 8.2:

Following the same notation as [43], refer also to [27] and [8], the proof of this lemma is based on the blocking approach, which consists of breaking down a strictly stationary sequence $(X_1, \ldots, X_n)$, into $2\upsilon_n$, equal-sized blocks, that each one is of length $n - 2\upsilon_n a_n$ that is, for $1 \leqslant j \leqslant \upsilon_n$,

$$\begin{aligned} H_j &= \{i : 2(j-1)a_n + 1 \leqslant i \leqslant (2j-1)a_n\}, \\ T_j &= \{i : (2j-1)a_n + 1 \leqslant i \leqslant 2ja_n\}, \\ R &= \{i : (2\upsilon_n a_n + 1 \leqslant i \leqslant n\}. \end{aligned}$$

The values of υ_n, a_n are given in the following. Another important component in this proof is the sequence of independent blocks $(\xi_1, \ldots, \xi_n)$ satisfying

$$\mathcal{L}(\xi_1, \ldots, \xi_n) = \mathcal{L}(X_1, \ldots, X_{a_n}) \times \mathcal{L}(X_{a_n+1}, \ldots, X_{2a_n}) \times \cdots .$$

As in [27, 43], applying the results of [84] on β-mixing, gives us for any measurable set A

$$\Big|\mathbb{P}\left\{(\xi_1, \ldots, \xi_{a_n}, \xi_{2a_n+1}, \ldots, \xi_{3a_n}, \ldots, \xi_{2(\upsilon_n-1)a_n+1}, \ldots, \xi_{2\upsilon_n a_n}) \in A\right\}$$

$$-\mathbb{P}\left\{(X_1,\ldots,X_{a_n},X_{2a_n+1},\ldots,X_{3a_n},\ldots,X_{2(\upsilon_n-1)a_n+1},\ldots,X_{2\upsilon_n a_n})\in A\right\}\Big|$$
$$\leqslant\ 2(\upsilon_n-1)\beta_{a_n}. \tag{A.1}$$

Let $t_1,\ldots,t_{N_\varepsilon(S_\mathcal{X})}$ be an ε-net of $S_\mathcal{X}$ and $\ell(t)=\arg\min_{\ell\in\{1,2,\ldots,N_\varepsilon(S_\mathcal{X})\}} d(t,t_\ell)$. Observing the following decomposition

$$\begin{aligned}
&\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le h_{n,2}}\ \sup_{t\in S_\mathcal{X}}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t,h_K(t))-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t,h_K(t))\right)\right|\\
&\le \underbrace{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le h_{n,2}}\ \sup_{t\in S_\mathcal{X}}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t,h_K(t))-\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t_\ell))\right|}_{G_1}\\
&+\underbrace{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le h_{n,2}}\ \sup_{t\in S_\mathcal{X}}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t_\ell))-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t_\ell))\right)\right|}_{G_2}\\
&+\underbrace{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le h_{n,2}}\ \sup_{t\in S_\mathcal{X}}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t_\ell))\right)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t,h_K(t))\right)\right|}_{G_3}.
\end{aligned} \tag{A.2}$$

Let us start with the term G_2. We need to prove that there exists $\eta>0$ and $b_0>0$, such that

$$\sum_{n\ge1}\mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le b_0}\ \sup_{t\in S_\mathcal{X}}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t))-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t))\right)\right|>\eta\sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right\}<\infty.$$

The key strategy here is to work with independent block sequence $\left\{\xi_j=(\zeta_j,\varsigma_j)\right\}_{j=1}^{\infty}$ instead of dependent once, this can be achieved by (A.1). Then we can clearly see that for $1\le\ell\le N_\varepsilon(S_\mathcal{X})$:

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le b_0}\ \sup_{t\in S_\mathcal{X}}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t))-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t))\right)\right|>\eta\sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right\}\\
&=\ \mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le b_0}\ \sup_{t\in S_\mathcal{X}}\left|(n\phi_t(h_K))^{-1}\sum_{i=1}^{n}\left(\varphi(Y_i)K\left(\frac{d(X_i,t_\ell)}{h_K}\right)-\mathbb{E}\left[\varphi(Y_i)K\left(\frac{d(X_i,t_\ell)}{h_K}\right)\right]\right)\right|\right.\\
&\qquad\left.>\eta\sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right\}\\
&\le\ 2\mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\le h_K\le b_0}\ \sup_{t\in S_\mathcal{X}}\left|(n\phi_t(h_K))^{-1}\sum_{j=1}^{\upsilon_n}\sum_{i\in H_j}\left(\varphi(\zeta_i)K\left(\frac{d(\varsigma_i,t)}{h_K}\right)-\mathbb{E}\left[\varphi(\zeta_i)K\left(\frac{d(\varsigma_i,t)}{h_K}\right)\right]\right)\right|\right.
\end{aligned}$$

$$> \eta' \sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\Bigg\} + 2(\upsilon_n - 1)\beta_{a_n}. \tag{A.3}$$

If we put

$$a_n = [(\log n)^{-1}(n^{p-2}\phi^p(h_K))^{1/2(p-1)}] \text{ and } \upsilon_n = \left[\frac{n}{2a_n}\right] - 1,$$

then, by condition **(C.5.1)**, it follows that $(\upsilon_n - 1)\beta_{a_n} \longrightarrow 0$ as $n \to 0$. Hence, only the first term in the right-hand side of (A.3) remains to be dealt with. Set

$$h_{K,j} = 2^j h_{n,1}, \quad L(n) = \max\left\{j : h_{K,j} \leq 2b_0\right\} \text{ and } \mathcal{H}_{K,j} := [h_{K,j-1}, h_{K,j}].$$

Then, we have $[h_{n,1}, b_0] \subseteq \bigcup_{j=1}^{L(n)} \mathcal{H}_{K,j}$. The empirical measure on $\{\alpha_n(X_i, Y_i) : i = 1, 2, \ldots, n\}$ is defined by:

$$\alpha_n(g) = \frac{1}{\sqrt{n}} \sum_{i}^{n} \left\{g\left(X_j, Y_j\right) - \mathbb{E}\left[g\left(X_j, Y_j\right)\right]\right\}.$$

For the original sequence, we write

$$Y_{j,g}(X_{a_n}) = \sum_{i\in H_j} \left\{g\left(X_j, Y_j\right) - \mathbb{E}\left[g\left(X_j, Y_j\right)\right]\right\}.$$

For the constructed independent block sequence $\boldsymbol{\xi}$, define

$$Z_{j,g}(\xi_{a_n}) = \sum_{i\in H_J} \{g(\xi_i) - \mathbb{E}[g(\xi_i)]\} \quad \text{and} \quad \tilde{\alpha}_{\upsilon_n}(g) = \frac{1}{\sqrt{n}} \sum_{1}^{\upsilon_n} Z_{j,g}.$$

Hence, we have

$$\widehat{r}_{n,2}^{(1)}(\varphi, t_{\ell(t)}, h_K(t)) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t_{\ell(t)}, h_K(t))\right) = \frac{1}{\sqrt{n}\phi_{t_{\ell(t)}}(h_K)}\tilde{\alpha}_{\upsilon_n}(\varphi K).$$

We consider also the following class of functions for $1 \leq j \leq L(n)$ and $1 \leq \ell \leq N_{\varepsilon_n}(S_{\mathcal{X}})$:

$$\mathscr{G}_{\varphi K,j}^{(\ell)} = \left\{(x, y) \mapsto \varphi(y)K\left(\frac{d(x, t_\ell)}{h_K}\right) \text{ where } K \in \mathscr{K}, \varphi \in \mathscr{F} \text{ and } h_K \in \mathcal{H}_{K,j}\right\},$$

with the envelope function $G_{\varphi K,j}^{(\ell)}$, also keep in mind that $S_{\mathcal{X}} \subseteq \bigcup_{\ell=1}^{N_\varepsilon(S_{\mathcal{X}})} B(t_\ell, \varepsilon)$. Now that we have all the necessary elements and using the fact that we are now dealing with sums of independent blocks, we can use the standard inequalities of independent setting. We have

$$\mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}\mathscr{K}} \sup_{h_{n,1}\leq h_K\leq b_0} \sup_{t\in S_{\mathcal{X}}} \left|(\sqrt{n}\phi_t(h_K))^{-1}\tilde{\alpha}_{\upsilon_n}(\varphi K)\right| > \eta' \sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right\}$$

$$\leq \sum_{\ell=1}^{N_{\varepsilon_n}(S_{\mathcal{X}})} \sum_{j=1}^{L(n)} \mathbb{P}\left\{ \frac{1}{\sqrt{n\phi(h_{K,j})\psi_{S_{\mathcal{X}}}(\varepsilon_n)}} \max_{1\leq \ell\leq N_{\varepsilon_n}(S_{\mathcal{X}})} \max_{1\leq j\leq L(n)} \left\| \sqrt{n}\tilde{\alpha}_{\upsilon_n}(K)\right\|_{\mathscr{G}^{(\ell)}_{K,j}} \geq \eta' \right\}$$

$$\leq L(n)N_{\varepsilon_n}(S_{\mathcal{X}}) \max_{1\leq j\leq L(n)} \max_{1\leq \ell\leq N_{\varepsilon_n}(S_{\mathcal{X}})} \mathbb{P}\left\{ \left\| \sqrt{n}\tilde{\alpha}_{\upsilon_n}(\varphi K)\right\|_{\mathscr{G}^{(\ell)}_{\varphi K,j}} > \eta' \sqrt{n\phi(h_{K,j})\psi_{S_{\mathcal{X}}}(\varepsilon_n)} \right\}$$

$$\leq L(n)N_{\varepsilon_n}(S_{\mathcal{X}}) \max_{1\leq j\leq L(n)} \max_{1\leq \ell\leq N_{\varepsilon_n}(S_{\mathcal{X}})} \mathbb{P}\left\{ \max_{1\leq p\leq n} \left\| \sqrt{p}\tilde{\alpha}_{\upsilon_p}(\varphi K)\right\|_{\mathscr{G}^{(\ell)}_{\varphi K,j}} > \eta' \sqrt{n\phi(h_{K,j})\psi_{S_{\mathcal{X}}}(\varepsilon_n)} \right\}. \quad \text{(A.4)}$$

Now, to bound the probability in (A.4), we apply Bernstein's inequality with

$$\mu_n := \mathbb{E}\left\| \sqrt{n}\alpha_{\upsilon_n}(\varphi K)\right\|_{\mathscr{G}^{(\ell)}_{\varphi K,j}},$$

and

$$\sigma^2 := \mathbb{E}\left(G^{(\ell)}_{\varphi K,j}(X,Y)\right)^2.$$

We know that the envelope function $G^{(\ell)}_{\varphi K,j}$, verifies

$$G^{(\ell)}_{\varphi K,j}(X,Y) \leq M\kappa_2 \mathbb{1}_{B(t_\ell,h_{K,j})}(X).$$

Hence, by **(C.1.1)**, **(C.3.1)**, **(C.4.1)** combined with Lemma B.1 and Lemma B.2, we get that

$$\mu_n = O\left(\sqrt{n\phi(h_{K,j})}\right) \text{ and } \sigma^2 = O\left(\phi(h_{K,j})\right).$$

Now, we can apply Bernstein's inequality to the empirical processes, for

$$z = (\eta'/2)\sqrt{n\phi(h_{K,j})\psi_{S_{\mathcal{X}}}(\varepsilon_n)} \text{ and } H = \kappa_2\left(\frac{2}{q!}\right)^{1/(q-2)}, \text{ or simply } H = \kappa_2 \text{ if } q = 2.$$

Making use of Lemma B.3, we get

$$\begin{aligned} &\mathbb{P}\left\{ \max_{1\leq p\leq n} \left\| \sqrt{p}\tilde{\alpha}_{\upsilon_p}(\varphi K)\right\|_{\mathscr{G}^{(\ell)}_{K,j}} \geq \eta' \sqrt{n\phi(h_{K,j})\psi_{S_{\mathcal{X}}}(\varepsilon_n)} \right\} \\ &\leq \mathbb{P}\left\{ \max_{1\leq p\leq n} \left\| \sqrt{p}\tilde{\alpha}_{\upsilon_p}(\varphi K)\right\|_{\mathscr{G}^{(\ell)}_{K,j}} \geq \mu_n + z \right\} \quad \text{(A.5)} \\ &\leq \exp\left(\frac{-(\eta'/2)^2\sqrt{n}\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{8\kappa_2^2\sqrt{n} + 4\eta'\kappa_2\sqrt{\frac{\psi_{S_{\mathcal{X}}}(\varepsilon_n)}{\phi(h_{K,j})}}} \right) \\ &\leq n^{-\mathbf{C}\epsilon_0}, \quad \text{(A.6)} \end{aligned}$$

where $\mathbf{C} > 0$. Moreover, from Eq (2.10), and the fact that $L(n) \sim \frac{\log(b_0/h_{n,1})}{\log(2)}$ and by choosing

$$\epsilon_0^2\mathbf{C} > 1, \varepsilon_n = \log n/n,$$

we get from (A.6) and (A.4),

$$\sum_{n\geq 1} \mathbb{P}\left\{ G_2 > \eta \sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}} \right\} \leq \sum_{n\geq 1} CL(n)N_{\varepsilon_n}(S_{\mathcal{X}})n^{-\mathbf{C}\epsilon_0} < \infty.$$

Now, let us study the term G_1. By conditions **(C.4.1)** and **(C.3.1)**, for some constant $\mathbb{C} > 0$, we have

$$\begin{aligned}
&\sup_{\varphi K\in\mathscr{F}\mathscr{K}} \sup_{h_{n,1}\le h_K\le b_0} \sup_{t\in S_{\mathcal{X}}} \left|\widehat{r}_{n,2}^{(1)}(\varphi,t,h_K(t)) - \widehat{r}_{n,2}^{(1)}(\varphi,t_{\ell(t)},h_K(t_\ell))\right| \\
&= \sup_{\varphi K\in\mathscr{F}\mathscr{K}} \sup_{h_{n,1}\le h_K\le b_0} \sup_{t\in S_{\mathcal{X}}} \left|\frac{1}{n\phi_t(h_K)}\sum_{i=1}^n \varphi(Y_i)K\left(\frac{d(X_i,t)}{h_K}\right) - \frac{1}{n\phi_{t_{\ell(t)}}(h_K)}\sum_{i=1}^n \varphi(Y_i)K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right| \\
&\le M\left|\frac{1}{n\phi_t(h_K)}\sum_{i=1}^n K\left(\frac{d(X_i,t)}{h_K}\right) - \frac{1}{n\phi_{t_{\ell(t)}}(h_K)}\sum_{i=1}^n K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right| \\
&\le \sup_{K\in\mathscr{K}} \sup_{h_{n,1}\le h_K\le b_0} \sup_{t\in S_{\mathcal{X}}} \frac{M\mathbb{C}}{n\phi(h_K)}\sum_{i=1}^n \left|K\left(\frac{d(X_i,t)}{h_K}\right) - K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right| \mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i).
\end{aligned}$$

Using the fact that the kernel $K(\cdot)$ is supposedly Lipschitz, we obtain

$$\begin{aligned}
G_1 &\le \sup_{h_{n,1}\le h_K\le b_0} \sup_{t\in S_{\mathcal{X}}} \frac{1}{n}\sum_{i=1}^n \frac{1}{h_K\phi(h_K)} d(t,t_{\ell(t)})\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i) \\
&\le \sup_{h_{n,1}\le h_K\le b_0} \sup_{t\in S_{\mathcal{X}}} \frac{1}{n}\sum_{i=1}^n \frac{\varepsilon_n}{h_K\phi(h_K)} \mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i) \\
&\le \sup_{h_{n,1}\le h_K\le b_0} \sup_{t\in S_{\mathcal{X}}} \frac{1}{n}\sum_{i=1}^n W_i,
\end{aligned}$$

where for $1 \le i \le n$:

$$W_i = \frac{\varepsilon_n}{h_K\phi(h_K)} \mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i).$$

Taking into account that $h_K \in [h_{n,1}, h_{n,2}]$ and using condition **(C.7.)**, we get for $q \ge 2$:

$$\mathbb{E}\,(W_1)^q \le C_1^q \left(\frac{1}{\sqrt{\phi(h_{n,1})}}\right)^{2(q-1)},$$

where we assume that C_1 is the bound of the sequence $\log n/nh_{n,1}$ which converges to 0. Hence, by applying a standard inequality (see Corollary A.8 [91]), the conditions are satisfied here, uniformly on $t \in S_{\mathcal{X}}$ and on h_K, on :

$$\frac{1}{n}\sum_{i=1}^n W_i,$$

in combination with condition **(C.6.)** and **(C.7.)**, we readily obtain

$$G_1 = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right). \tag{A.7}$$

Finally, we treat G_3. Noticing that

$$G_3 \le \mathbb{E}\left(\sup_{\varphi K\in\mathscr{F}\mathscr{K}} \sup_{h_{n,1}\le h_K\le b_0} \sup_{t\in S_{\mathcal{X}}} \left|\widehat{r}_{n,2}^{(1)}(\varphi,t,h_K(t)) - \widehat{r}_{n,2}^{(1)}(\varphi,t_\ell,h_K(t_\ell))\right|\right),$$

similar to the procedures of treating G_1 once again, it follows directly that

$$G_3 = O_{a.co}\left(\sqrt{\frac{\psi_{S_X}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right). \tag{A.8}$$

Hence, by combining (A.2) and (A.8), we so obtain (8.2), as sought.

Proof of (8.3)*:*

Notice that (8.3) is a direct result of (8.2) when the function $\varphi \equiv 1$. This completes the proof of the lemma.

Proof of Lemma 8.10

Let us consider the term $I_1\,(t, H_{n,k}(t))$ such that $\mathbb{1}_{\{D_n^- \le H_{n,k}(t) \le D_n^+\}} \xrightarrow{a.co} 1$ when $\frac{k}{n} \to 0$, and making use of condition **(C.3.2)**, we have

$$\widehat{r}_{n,2}^{(1)}(\varphi, t; D_n^+) \le \widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t)) \le \widehat{r}_{n,2}^{(1)}(\varphi, t; D_n^-). \tag{A.9}$$

We have

$$\begin{aligned}
&\sqrt{k}\left|S_1\,(t, H_{n,k}(t)) - \mathbb{E}\,(S_1\,(t, H_{n,k}(t)))\right| \\
&\le \quad \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t)) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t))\right)\right| && \text{(A.10)}\\
&\quad + \sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t))\right) - \widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t))\right| && \text{(A.11)}\\
&= \quad I_{11}\,(t, H_{n,k}(t)) + I_{12}\,(t, H_{n,k}(t)).
\end{aligned}$$

The Eq (A.10) gives

$$\begin{aligned}
\left|I_{11}\,(t, H_{n,k}(t))\right| &= \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t)) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t))\right)\right| \\
&\le \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^-) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^+)\right)\right| \\
&\le \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^-) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^-)\right)\right| && \text{(A.12)}\\
&\quad + \sqrt{k}\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^-)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^+)\right). && \text{(A.13)}
\end{aligned}$$

Similarly, (A.11) gives us:

$$\begin{aligned}
\left|I_{12}\,(t, H_{n,k}(t))\right| &= \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t)) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t; H_{n,k}(t))\right)\right| \\
&\le \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^-) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^-)\right)\right| \\
&\quad + \sqrt{k}\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^-)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^+)\right). && \text{(A.14)}
\end{aligned}$$

The following decomposition will be used to treat (A.12)

$$
\begin{aligned}
&\sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t,D_n^-)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t,D_n^-)\right)\right| \\
&\leq \quad \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t,D_n^-)-\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right| && \text{(A.15)}\\
&\quad + \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right)\right| && \text{(A.16)}\\
&\quad + \sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t,D_n^-)\right)\right|. && \text{(A.17)}
\end{aligned}
$$

For some $\xi_n \in (0,1)$, recall that D_n^- and D_n^+ are defined in (8.6) and (8.7) respectively. Moreover, observe that

$$
\begin{aligned}
&\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t,D_n^-)\right)\right| \\
&= \quad \frac{1}{n\mathbb{E}\left(\Delta_1\left(t,D_n^-\right)\right)}\left|\mathbb{E}\left(\sum_{i=1}^{n}\varphi(Y_i)\mathbb{1}_{\{|F(Y_i)|>\delta_n\}}\Delta_i\left(t,D_n^-\right)\right)\right| \\
&\leq \quad \mathbb{E}\left(|\varphi(Y)|\,\mathbb{1}_{\{|F(Y)|>\delta_n\}}\Delta_1\left(t,D_n^-\right)\right)\left(\mathbb{E}\left(\Delta_1\left(t,D_n^-\right)\right)\right)^{-1}.
\end{aligned}
$$

Under the assumptions **(C.4.2)** and **(C.8.)**, using Hölder's inequality, for $\alpha_1 = \frac{p}{2}$ and α_2 such that $\frac{1}{\alpha_1}+\frac{1}{\alpha_2}=1$, we can write

$$
\begin{aligned}
&\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi,t,D_n^-)\right)\right| \\
&\leq \quad \frac{\sqrt{k}}{\phi\left(D_n^-\right)}\mathbb{E}^{1/\alpha_1}\left[\left|F(Y)_1^{\alpha}\right|\mathbb{1}_{\{|F(Y)|>\delta_n\}}\right]\mathbb{E}^{1/\alpha_2}\left[\Delta_1^{\alpha_2}\left(t,D_n^-\right)\right] \\
&\leq \quad \frac{\sqrt{k}}{\phi\left(D_n^-\right)}\delta_n^{-1}\mathbb{E}^{1/\alpha_1}\left(|F(Y)|^p \mid X\right)\phi^{1/\alpha_2}\left(D_n^-\right) \\
&\leq \quad \frac{C\sqrt{k}}{\delta_n\phi^{2/p}\left(D_n^-\right)} \\
&\leq \quad \frac{C}{\zeta^{\frac{1}{p}}\sqrt{n}\delta_n}\left(\frac{k}{n}\right)^{\frac{1}{2}-\frac{2}{p}} \to 0 \text{ as } n\to\infty.
\end{aligned}
$$

To obtain the convergence in probability of (A.15), it suffices to use the assumption **(C.4.2)**, **(C.8.)**, and Markov's inequality. Indeed, for all $\varepsilon>0$, we obtain

$$
\begin{aligned}
&\mathbb{P}\left(\sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi,t,D_n^-)-\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right| > \sqrt{k}\varepsilon\right) \\
&\leq \quad \mathbb{E}\left(|F(Y)|\,\mathbb{1}_{\{|F(Y)|>\delta_n\}}\Delta_1\left(t,D_n^-\right)\right)\left(\sqrt{k}\varepsilon\mathbb{E}\left(\Delta_1\left(t,D_n^-\right)\right)\right)^{-1} \\
&\leq \quad \frac{C\phi^{(1-\alpha_2)/\alpha_2}\left(D_n^-\right)}{\varepsilon\delta_n\sqrt{k}} \\
&\leq \quad \frac{C}{\varepsilon\delta_n}\frac{\phi^{-\frac{2}{p}}\left(D_n^-\right)}{\sqrt{k}} \to 0
\end{aligned}
$$

for some δ_n large enough and by condition **(C.8.)**. Now, for the second term of (A.14), Lemma 8.2 gives

$$\left|\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^-) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^-)\right)\right| \to 0, \ a.co. \ \text{as} \ n \to \infty. \tag{A.18}$$

Using the fact that $k\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^{2\gamma} \to 0$ and by condition **(C.2.1)**, for $u = u_n = \phi^{-1}\left(\frac{k}{n}\right)$, (A.13) gives

$$\begin{aligned}
&\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^-)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^+)\right)\right| \\
&\leq \sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^-)\right) - r^{(1)}(\varphi, t)\right| \\
&\quad + \sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^+)\right) - r^{(1)}(\varphi, t)\right| \\
&\leq \sqrt{k}\left|\mathbb{E}\left(\frac{1}{\mathbb{E}\Delta_1\left(t, D_n^-\right)}\Delta_1\left(t, D_n^-\right)\varphi(Y)\right) - r^{(1)}(\varphi, t)\right| \\
&\quad + \sqrt{k}\left|\mathbb{E}\left(\frac{1}{\mathbb{E}\Delta_1\left(t, D_n^+\right)}\Delta_1\left(t, D_n^+\right)\varphi(Y)\right) - r^{(1)}(\varphi, t)\right| \\
&\leq \sqrt{k}\sum_{u=D_n^-, D_n^+}\left|\frac{1}{\mathbb{E}\Delta_1(t, u)}\mathbb{E}'\left(\Delta_1(t, u)\mathbb{1}_{B(t,u)}\left(r^{(1)}\left(\varphi, X_1\right) - r^{(1)}(\varphi, t)\right)\right)\right| \\
&\leq \sqrt{k}\left(C\mathbb{1}_{B\left(t, D_n^-\right)} + C'\mathbb{1}_{B\left(t, D_n^+\right)}\right) d(t, X_1)^\gamma \\
&\leq C\sqrt{k}\left((D_n^-)^\gamma + (D_n^+)^\gamma\right) \\
&\leq \frac{2C}{\zeta^\gamma}\sqrt{k\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^{2\gamma}} \to 0, \ \text{as} \ n \to \infty.
\end{aligned} \tag{A.19}$$

On the other side, recall that $\varphi_M(Y) := \varphi(Y)\mathbb{1}_{\{F(Y)<M\}}$, then, using the fact that the regression function satisfies the Lipschitz condition and under condition **(C.4.1)**, we have

$$\begin{aligned}
&\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^-)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^+)\right)\right| \\
&\leq \sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^-)\right) - r^{(1)}(\varphi_M, t)\right| \\
&\quad + \left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M, t, D_n^+)\right) - r^{(1)}(\varphi_M, t)\right| \\
&\leq \sqrt{k}\left|\mathbb{E}\left(\frac{1}{\mathbb{E}\Delta_1\left(t, D_n^-\right)}\Delta_1\left(t, D_n^-\right)\varphi_M(Y)\right) - r^{(1)}(\varphi_M, t)\right| \\
&\quad + \sqrt{k}\left|\mathbb{E}\left(\frac{1}{\mathbb{E}\Delta_1\left(t, D_n^+\right)}\Delta_1\left(t, D_n^+\right)\varphi_M(Y)\right) - r^{(1)}(\varphi_M, t)\right| \\
&\leq \sqrt{k}\sum_{u=D_n^-, D_n^+}\left|\frac{1}{\mathbb{E}\Delta_1(t, u)}\mathbb{E}\left(\Delta_1(t, u)\mathbb{1}_{B(t,u)}\left(r^{(1)}(\varphi_M, X_1) - r^{(1)}(\varphi_M, t)\right)\right)\right| \\
&\leq \sqrt{k}\left(C\mathbb{1}_{B\left(t, D_n^-\right)} + C'\mathbb{1}_{B\left(t, D_n^+\right)}\right) d(t, X_1)^\gamma \\
&\leq \frac{2C}{\zeta^\gamma}\sqrt{k\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^{2\gamma}} \to 0, \ \text{as} \ n \to \infty.
\end{aligned} \tag{A.20}$$

For the second term of the right-hand side of the Eq (8.74), we have

$$\left|S_2(t, H_{n,k}(t))\right| = \left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t; H_{n,k}(t)) - r^{(1)}(\varphi, t)\right)\right| \tag{A.21}$$
$$\leq \left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi, t, D_n^-) - r^{(1)}(\varphi, t)\right)\right|, \tag{A.22}$$

making use of the Lemma 8.2, we readily infer

$$\sqrt{k}S_2(t, H_{n,k}(t)) = O\left(\sqrt{\zeta^\gamma k\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^{2\gamma}}\right). \tag{A.23}$$

The proof of $S_2(t, H_{n,k}(t))$ is complete. Hence, the convergence of $I_1(t, H_{n,k}(t))$ to 0 is complete combining the result (8.74) and (A.18), (A.19), (A.20) and (A.23).

Proof of Lemma 8.3:

In this lemma, we suppose that the class of functions $\mathscr{F}$ is not necessarily bounded but satisfying condition **(C.4.2)**. In order to prove Lemma 8.3, we proceed as follows, for an arbitrary $\lambda > 0$ and each function $\varphi \in \mathscr{F}$:

$$\begin{aligned}\varphi(y) &= \varphi(y)\mathbb{1}_{\left\{F(y)\leq\lambda\eta_n^{1/p}\right\}} + \varphi(y)\mathbb{1}_{\left\{F(y)>\lambda\eta_n^{1/p}\right\}}\\ &= \varphi^{(T)}(y) + \varphi^{(R)}(y),\end{aligned}$$

where we take $\eta_n = n/(\log n)^2$. So, we write

$$\begin{aligned}\widehat{r}_{n,2}^{(1)}(\varphi, t, h_K(t)) &:= \frac{1}{n\mathbb{E}\left(\Delta_1(t; h_K(t))\right)}\sum_{i=1}^n \varphi^{(T)}(Y_i)K\left(\frac{d(X_i, t)}{h_K}\right)\\ &\quad+\frac{1}{n\mathbb{E}\left(\Delta_1(t; h_K(t))\right)}\sum_{i=1}^n \varphi^{(R)}(Y_i)K\left(\frac{d(X_i, t)}{h_K}\right)\\ &= \widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t)) + \widehat{r}_{n,2}^{(1)}(\varphi^{(R)}, t, h_K(t)).\end{aligned}$$

Let us start with the truncated part. Following the same reasoning as in the proof of Lemma 8.2, we have:

$$\begin{aligned}&\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\leq h_K\leq h_{n,2}}\ \sup_{t\in S_{\mathcal{X}}}\left|\widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t)) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t))\right)\right|\\ &\leq \underbrace{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\leq h_K\leq h_{n,2}}\ \sup_{t\in S_{\mathcal{X}}}\left|\widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t)) - \widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t_\ell)\right)\right|}_{G_{1,T}}\\ &\quad+\underbrace{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\leq h_K\leq h_{n,2}}\ \sup_{t\in S_{\mathcal{X}}}\left|\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t_\ell)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t_\ell)\right)\right)\right|}_{G_{2,T}}\\ &\quad+\underbrace{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\ \sup_{h_{n,1}\leq h_K\leq h_{n,2}}\ \sup_{t\in S_{\mathcal{X}}}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t_\ell)\right)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t))\right)\right|}_{G_{3,T}}.\end{aligned} \tag{A.24}$$

First, let us start with the term $G_{2,T}$, using the same notation as the proof of Lemma 8.2, we have:

$$\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t)\right)\right) = \frac{1}{\sqrt{n}\phi_{t_{\ell(t)}}(h_K)}\tilde{\alpha}_{\upsilon_n}(\varphi^{(T)}K).$$

where $\tilde{\alpha}_{\upsilon_n}(\varphi^{(T)}K)$ is an empirical process indexed by the class of functions $\mathscr{F}\mathscr{K}\mathscr{I}_{\lambda\eta_n}$, where we define the class of functions of $y \in \mathcal{Y}$, $\mathscr{I}_{\lambda\eta_n}$ by

$$\mathscr{I}_{\lambda\eta_n} := \left\{\mathbb{1}_{\left\{F(y)\leq\lambda\eta_n^{1/p}\right\}} : \lambda > 0\right\},$$

where p being the moment order given in **(C.4.2)**. We consider also the following class of functions, for $1 \leq j \leq L(n)$ and $1 \leq \ell \leq N_{\varepsilon_n}(S_{\mathcal{X}})$,

$$\mathscr{G}_{\varphi K,j}^{(\ell,T)} = \left\{(x,y) \mapsto \varphi(y)K\left(\frac{d(x,t_\ell)}{h_K}\right)\mathbb{1}_{\left\{F(y)\leq\lambda\eta_n^{1/p}\right\}} : K \in \mathscr{K}, \varphi \in \mathscr{F} \text{ and } h_K \in \mathcal{H}_{K,j}, \lambda > 0\right\},$$

that the envelope function is denoted by $G_{\varphi K,j}^{(\ell,T)}$, then we have

$$\mathbb{P}\left\{\sup_{\varphi K\in\mathscr{F}\mathscr{K}}\sup_{h_{n,1}\leq h_K\leq b_0}\sup_{t\in S_{\mathcal{X}}}\left|\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t)\right)\right)\right| > \eta\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right\} \quad (A.25)$$

$$\leq L(n)N_{\varepsilon_n}(S_{\mathcal{X}})\max_{1\leq j\leq L(n)}\max_{1\leq\ell\leq N_{\varepsilon_n}(S_{\mathcal{X}})}\mathbb{P}\left\{\max_{1\leq p\leq n}\left\|\sqrt{p}\tilde{\alpha}_{\upsilon_p}(\varphi^{(T)}K)\right\|_{\mathscr{G}_{\varphi K,j}^{(\ell)}} > \eta'\sqrt{n\phi(h_{K,j})\psi_{S_{\mathcal{X}}}(\varepsilon_n)}\right\}. \quad (A.26)$$

We apply Bernstein's inequality to bound (A.26), so, we have to study the asymptotic behavior of the quantities

$$\mu_{n,T}^{(2)} = \mathbb{E}\left\|\sqrt{n}\tilde{\alpha}_{\upsilon_n}(\varphi^{(T)}K)\right\|_{\mathscr{G}_{\varphi K,j}^{(\ell,T)}},$$

and

$$\mathbb{E}\left(G_{\varphi K,j}^{(\ell,T)}(X,Y)\right)^2.$$

From condition **(C.4.2)**, we remark that

$$G_{\varphi K,j}^{(\ell,T)}(X,Y) \leq F(Y)\kappa_2\mathbb{1}_{B(t_\ell,h_{K,j})}(X)\mathbb{1}_{\left\{F(y)\leq\lambda\eta_n^{1/p}\right\}},$$

so thanks to the condition **(C.4.2)**, we get

$$\mathbb{E}\left(G_{\varphi K,j}^{(\ell,T)}(X,Y)\right)^2 \leq \lambda^2\kappa_2^2\phi(h_{K,j})\eta_n^{2/p} = \sigma_{(T)}^2.$$

Furthermore, making use of Lemma B.1 in combination with Lemma B.2 in appendix, and since condition **(C.4.3)** is satisfied, we infer that

$$\mu_{n,T}^{(2)} = \mathbb{E}\left\|\sqrt{n}\tilde{\alpha}_{\upsilon_n}(\varphi^{(T)}K)\right\|_{\mathscr{G}_{\varphi K,j}^{(\ell,T)}} \leq C'\kappa_2\sqrt{n\phi(h_{K,j})\eta_n^{2/p}},$$

which means that,

$$\mu_{n,T}^{(2)} = O\left(\sqrt{n\phi(h_{K,j})\eta_n^{2/p}}\right).$$

Now, we can apply Bernstein's inequality for empirical processes with:

$$z = \eta' \sqrt{n\phi(h_{K,j})\psi_{S_\mathcal{X}}(\varepsilon_n)},$$

in Lemma B.3, we obtain, for $\mathfrak{C}' > 0$,

$$\begin{aligned}
&\mathbb{P}\left\{\max_{1\le k\le n}\left\|\tilde{\alpha}_{\upsilon_k}(\varphi^{(T)}K)\right\|_{\mathscr{G}_{\varphi K,j}^{(\ell,T)}} \geq \eta' \sqrt{n\phi(h_{K,j})\psi_{S_\mathcal{X}}(\varepsilon_n)}\right\}\\
&\leq \mathbb{P}\left\{\max_{1\le k\le n}\left\|\sqrt{k}\tilde{\alpha}_{\upsilon_k}(\varphi^{(T)}K)\right\|_{\mathscr{G}_{\varphi K,j}^{(\ell,T)}} \geq \mu_{n,T}^{(2)} + z\right\}\\
&\leq \exp\left(-\eta'^2 \frac{n\phi(h_{K,j})\psi_{S_\mathcal{X}}(\varepsilon_n)}{2n\phi(h_{K,j})\eta_n^{2/p} + \mathfrak{C}'\sqrt{n\phi(h_{K,j})\psi_{S_\mathcal{X}}(\varepsilon_n)}}\right)\\
&\leq n^{-\eta_0'^2 \mathbf{C}'},
\end{aligned} \tag{A.27}$$

where $\mathbf{C}' > 0$. From Eq (2.10), and fact that $L(n) \sim \dfrac{\log(b_0/h_{n,1})}{\log(2)}$ and by choosing

$$\varepsilon_n = \frac{\log n}{n},$$

we get

$$\sum_{n\geq 1}\mathbb{P}\left\{G_{2,T} > \eta\sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right\} \leq \sum_{n\geq 1} CL(n)N_{\varepsilon_n}(S_\mathcal{X})n^{-\mathbf{C}'\eta_0'} < \infty.$$

Next, we prove that the result is valid for the term $G_{1,T}$. By conditions **(C.4.2)** and **(C.3.1)**, for some constant $\mathbb{C}' > 0$, we have

$$\begin{aligned}
&\left|\widehat{r}_{n,2}^{(1)}(\varphi, t, h_K(t)) - \widehat{r}_{n,2}^{(1)}(\varphi, t_{\ell(t)}, h_K(t_\ell))\right|\\
&= \left|\frac{1}{n\phi_t(h_K)}\sum_{i=1}^n \varphi^{(T)}(Y_i)K\left(\frac{d(X_i,t)}{h_K}\right) - \frac{1}{n\phi_{t_{\ell(t)}}(h_K)}\sum_{i=1}^n \varphi^{(T)}(Y_i)K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right|\\
&\leq \frac{\mathbb{C}'}{n\phi(h_K)}\left|\sum_{i=1}^n F^{(T)}(Y_i)K\left(\frac{d(X_i,t)}{h_K}\right) - \sum_{i=1}^n F^{(T)}(Y_i)K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right|,
\end{aligned}$$

where

$$F^{(T)}(y) := F(y)\mathbb{1}_{\left\{F(y)\leq \lambda\eta_n^{1/p}\right\}}.$$

Observe that

$$\begin{aligned}
&\sup_{\varphi^{(T)}K\in\mathscr{F}\mathscr{K}\mathscr{I}_{\lambda\eta_n}}\ \sup_{h_{n,1}\le h_K\le b_0}\ \sup_{t\in S_\mathcal{X}}\left|\widehat{r}^{(1)}_{n,2}(\varphi,t,h_K(t))-\widehat{r}^{(1)}_{n,2}\left(\varphi,t_{\ell(t)},h_K(t_\ell)\right)\right|\\
\le&\sup_{\varphi^{(T)}K\in\mathscr{F}\mathscr{K}\mathscr{I}_{\lambda\eta_n}}\ \sup_{t\in S_\mathcal{X}}\ \sup_{h_{n,1}\le h_K\le b_0}\frac{1}{n\phi(h_K)}\sum_{i=1}^n\left|F^{(T)}(Y_i)\right|\left|K\left(\frac{d(X_i,t)}{h_K}\right)-K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right|\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i)\\
:=&\left|\mathscr{G}_{1,T}\right|.
\end{aligned}$$

Furthermore, following the same steps as the proof of G_1 in the previous lemma's proof, we have:

$$\begin{aligned}
\left|\mathscr{G}_{1,T}\right| &\le \sup_{t\in S_\mathcal{X}}\ \sup_{h_{n,1}\le h_K\le b_0}\frac{1}{n}\sum_{i=1}^n\frac{1}{h_K\phi(h_K)}\left|F^{(T)}(Y_i)\right|d(t,t_{\ell(t)})\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i)\\
&\le \sup_{t\in S_\mathcal{X}}\ \sup_{h_{n,1}\le h_K\le b_0}\frac{1}{n}\sum_{i=1}^n\frac{\varepsilon_n}{h_K\phi(h_K)}\left|F^{(T)}(Y_i)\right|\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i)\\
&= \sup_{t\in S_\mathcal{X}}\ \sup_{h_{n,1}\le h_K\le b_0}\frac{1}{n}\sum_{i=1}^n W_{i,T},
\end{aligned}$$

where, for $1\le i\le n$,

$$W_{i,T}=\frac{\varepsilon_n}{h_K\phi(h_K)}\left|F^{(T)}(Y_i)\right|\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i).$$

We get uniformly on $t\in S_\mathcal{X}$ and on h_K :

$$\sup_{t\in S_\mathcal{X}}\ \sup_{h_{n,1}\le h_K\le b_0}W_{1,T}\ \le\ \sup_{t\in S_\mathcal{X}}\frac{\varepsilon_n}{h_{n,1}\phi(h_{n,1})}\left|F^{(T)}(Y_1)\right|.$$

So, for $2\le q\le p$, we obtain

$$\begin{aligned}
\mathbb{E}\left(W_{1,T}\right)^q &\le \left(\frac{\varepsilon_n}{h_{n,1}\phi(h_{n,1})}\right)^q\mathbb{E}\left(\left|F^{(T)}(Y_1)\right|^q\right)\\
&\le \sup_{t\in S_\mathcal{X}}\left(\frac{\varepsilon_n}{h_{n,1}\phi(h_{n,1})}\right)^q\mathbb{E}\left(\left|F^{(T)}(Y_1)\right|^q\left|X=t\right.\right) \qquad (A.28)\\
&\le \left(\frac{\varepsilon_n}{h_{n,1}\phi(h_{n,1})}\right)^q\theta_p^{q/p}=\frac{\theta_p^{q/p}\varepsilon_n^q}{h_{n,1}^q\phi(h_{n,1})}\left(\frac{1}{\phi(h_{n,1})}\right)^{q-1}. \qquad (A.29)
\end{aligned}$$

The transition from (A.28) to (A.29) is done by using Jensen's inequality used for the concave function z^a, for $0<a\le 1$. Moreover, from condition **(C.7.)**, we deduce that the quantity $\varepsilon_n/h_{n,1}\phi(h_{n,1})$ is bounded, then for $q\ge 2$

$$\sup_{t\in S_\mathcal{X}}\ \sup_{h_{n,1}\le h_K\le b_0}\mathbb{E}\left(W_{1,T}\right)^q\ \le\ \mathfrak{C}\theta_p^{q/p}\left(\frac{1}{\sqrt{\phi(h_{n,1})}}\right)^{2(q-1)},$$

where $\mathfrak{C}>0$. Hence, by applying a standard inequality (see Corollary A.8 [91]) that the conditions are satisfied here, uniformly on $t\in S_\mathcal{X}$, we get

$$G_{1,T}=O_{a.co}\left(\sqrt{\frac{\psi_{S_\mathcal{X}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right).$$

Finally, all that is left is to evaluate the term $G_{3,T}$. We have

$$\begin{aligned}
&\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t_\ell)\right)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t))\right)\right| \\
&= \left|\mathbb{E}\left[\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t_\ell)\right) - \widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t))\right]\right| \\
&\leq \frac{1}{n\phi(h_K)}\left|\sum_{i=1}^{n}\mathbb{E}\left(F^{(T)}(Y_i)K\left(\frac{d(X_i,t)}{h_K}\right) - F^{(T)}(Y_i)K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right)\right|,
\end{aligned}$$

which means that

$$\begin{aligned}
&\sup_{\varphi^{(T)}K\in\mathscr{F}\mathscr{K}\mathscr{I}_{\lambda\eta_n}}\sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}\left(\varphi^{(T)}, t_{\ell(t)}, h_K(t_\ell)\right)\right) - \mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi^{(T)}, t, h_K(t))\right)\right| \\
&\leq \sup_{\varphi^{(T)}K\in\mathscr{F}\mathscr{K}\mathscr{I}_{\lambda\eta_n}}\sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\frac{1}{n\phi(h_K)}\sum_{i=1}^{n}\mathbb{E}\Bigg(\left|F^{(T)}(Y_i)\right| \\
&\quad \left|K\left(\frac{d(X_i,t)}{h_K}\right) - K\left(\frac{d(X_i,t_{\ell(t)})}{h_K}\right)\right|\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i)\Bigg) \\
&=: \left|\mathscr{G}_{3,T}(\varphi K, t; h_K)\right|.
\end{aligned}$$

Let's take a look at $\left|\mathscr{G}_{1,T}(\varphi K, t; h_K)\right|$ and following the same steps, by the fact that $K(\cdot)$ is a Lipschitz function, we get

$$\begin{aligned}
\left|\mathscr{G}_{3,T}(\varphi K, t; h_K)\right| &\leq \sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_K\phi(h_K)}\mathbb{E}\left(\left|F^{(T)}(Y_i)\right| d(t,t_{\ell(t)})\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i)\right) \\
&\leq \sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\frac{1}{n}\sum_{i=1}^{n}\frac{\varepsilon_n}{h_K\phi(h_K)}\mathbb{E}\left(\left|F^{(T)}(Y_i)\right|\mathbb{1}_{B(t,h_K)\cup B(t_{\ell(t)},h_K)}(X_i)\right) \\
&\leq \sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\frac{1}{n}\sum_{i=1}^{n}\frac{\varepsilon_n}{h_K\phi(h_K)}\mathbb{E}\left(\left|F^{(T)}(Y_i)\right|\right) \\
&\leq \sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\frac{1}{n}\sum_{i=1}^{n}\frac{\varepsilon_n}{h_K\phi(h_K)}\mathbb{E}\left(\mathbb{E}\left(F^{(T)}(Y_i)\middle|X=t\right)\right) \\
&= \sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\frac{1}{n}\sum_{i=1}^{n}W_{i,T}^{(3)}, \qquad \text{(A.30)}
\end{aligned}$$

where, for $1\leq i\leq n$,

$$W_{i,T}^{(3)} = \frac{\varepsilon_n}{h_K\phi(h_K)}\mathbb{E}\left(\left|F^{(T)}(Y_i)\right|\right).$$

We get for $q\geq 2$ and uniformly on $t\in S_{\mathcal{X}}$ and on h_K :

$$\begin{aligned}
\sup_{t\in S_{\mathcal{X}}}\sup_{h_{n,1}\leq h_K\leq b_0}\mathbb{E}\left(Z_{1,T}^{(31)}\right)^q &\leq \sup_{t\in S_{\mathcal{X}}}\left(\frac{\varepsilon_n}{h_{n,1}\phi(h_{n,1})}\right)^q\mathbb{E}\left(\mathbb{E}\left(\left|F^{(T)}(Y_1)\right|^q\middle|X=t\right)\right) \qquad \text{(A.31)} \\
&\leq \left(\frac{\varepsilon_n}{h_{n,1}\phi(h_{n,1})}\right)^q\theta_p^{q/p} = \frac{\theta_p^{q/p}\varepsilon_n^q}{h_{n,1}^q\phi(h_{n,1})}\left(\frac{1}{\phi(h_{n,1})}\right)^{q-1}. \qquad \text{(A.32)}
\end{aligned}$$

The transition from (A.31) to (A.32) is done by using Minkowski's and Jensen's inequalities. Then, following the same way as in treating $\left|\mathscr{G}_{1,T}(\varphi K, t; h_K)\right|$, we get

$$\sup_{\varphi^{(T)} K\in\mathscr{F}\mathscr{K}\mathscr{I}_{\lambda\eta_n}} \sup_{t\in S_{\mathcal{X}}} \sup_{h_{n,1}\leq h_K\leq b_0} \left|\mathscr{G}_{3,T}(\varphi K, t; h_K)\right| = O_{a.co}\left(\sqrt{\frac{\psi_{S_{\mathcal{X}}}\left(\frac{\log n}{n}\right)}{n\phi(h_{n,1})}}\right).$$

Therefore, the proof is done for the truncated part. Regarding the remainder term, the idea consists of proving its asymptotic negligibility, that is

$$\sup_{h_{n,1}\leq h_K\leq b_n} \sup_{t\in S_{\mathcal{X}}} \sup_{\varphi K\in\mathscr{F}\mathscr{K}} \left|\widehat{r}^{(1)}_{n,2}(\varphi^{(R)}, t, h_K(t)) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi^{(R)}, t, h_K(t))\right)\right| = o_{a.co}(1),$$

which could be derived directly from the proof of the remainder part of the U-statistics developed in the sequel.

Proof of Lemma 8.11

We recall that

$$\begin{aligned} I_2(t,u) &= \widehat{r}^{(1)}_{n,2}(\varphi_M, t, u) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, u)\right) - r^{(1)}(\varphi, t)\left(\widehat{r}^{(1)}_{n,1}(1, t, u) - 1\right) \\ &= \left[\widehat{r}^{(1)}_{n,2}(\varphi_M, t, u) - r^{(1)}(\varphi, t)\widehat{r}^{(1)}_{n,1}(1, t, u)\right] - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, u) - r^{(1)}(\varphi, t)\widehat{r}^{(1)}_{n,1}(1, t, u)\right). \end{aligned}$$

Therefore, the proof of this lemma is based on the following decomposition:

$$\begin{aligned} &\sqrt{k}I_2(\varphi, t, H_{n,k}(t)) \\ &= \sqrt{k}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, H_{n,k}(t)) - r^{(1)}(\varphi, t)\widehat{r}^{(1)}_{n,1}(1, t, H_{n,k}(t)) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, H_{n,k}(t)) - r^{(1)}(\varphi, t)\widehat{r}^{(1)}_{n,1}(1, t, H_{n,k}(t))\right)\right) \\ &= \sqrt{k}\left[\widehat{r}^{(1)}_{n,2}(\varphi_M, t, H_{n,k}(t)) - \widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^+)\right] \\ &\quad + \sqrt{k}\left[\widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^+) - r^{(1)}(\varphi, t)\widehat{r}^{(1)}_{n,1}(1, t, D_n^+) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^+) - r^{(1)}(\varphi, t)\widehat{r}^{(1)}_{n,1}(1, t, D_n^+)\right)\right] \\ &\quad + \sqrt{k}\left[r^{(1)}(\varphi, t)\left(\widehat{r}^{(1)}_{n,1}(1, t, D_n^+) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,1}(1, t, D_n^+)\right)\right)\right] \\ &\quad + \sqrt{k}\left[r^{(1)}(\varphi, t)\left(\mathbb{E}\left(\widehat{r}^{(1)}_{n,1}(1, t, H_{n,k}(t))\right) - \widehat{r}^{(1)}_{n,1}(1, t, H_{n,k}(t))\right)\right] \\ &\quad + \sqrt{k}\left[\mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^+)\right) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, H_{n,k}(t))\right)\right] \\ &= J_1(x) + J_2(x) + J_3(x) + J_4(x) + J_5(x). \end{aligned} \tag{A.33}$$

By the same arguments as those involved in the proof of $I_1(\varphi, t, H_{n,k}(t))$. We get

$$\begin{aligned} |J_1(x)| &= \sqrt{k}\left|\widehat{r}^{(1)}_{n,2}(\varphi_M, t, H_{n,k}(t)) - \widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^+)\right| \\ &\leq \sqrt{k}\left|\widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^-) - \widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^+)\right| \\ &\leq \sqrt{k}\left|\widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^-) - \mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi_M, t, D_n^-)\right)\right| \end{aligned}$$

$$+\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right)-\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right|$$
$$+\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right)\right|, \quad (A.34)$$

and

$$\begin{aligned}|J_5(x)| &= \sqrt{k}\left(\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,H_{n,k}(t))\right)\right)\\ &\leq \sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right)-\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right|\\ &\quad+\sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,H_{n,k}(t))\right)\right| \quad (A.35)\\ &\leq 2\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right)-\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right|.\end{aligned}$$

By combining (A.34) and (A.35) and by similar arguments used to treat the $I_{11}(\varphi,t,H_{n,k}(t))$ and $I_{12}(\varphi,t,H_{n,k}(t))$, we get

$$\begin{aligned}|J_1(x)+J_5(x)| \leq\ & \sqrt{k}\left|\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right)\right|\\ &+3\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right)-\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right|\\ &+\sqrt{k}\left|\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^-)\right)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)\right)\right|.\end{aligned}$$

We then obtain that $|J_1(x)+J_5(x)|$ tending to zero as n tends to infinity. We evaluate the term $J_2(x)$ on the right side of (A.33). Let us introduce the following sum

$$J_2(x)=\sum_{i=1}^{n}\mathbb{Z}_{ni},$$

where

$$\mathbb{Z}_{ni}=\frac{\sqrt{k}}{\sqrt{n}\mathbb{E}\left(\Delta_1(t,D_n^+)\right)}\left(\left(\varphi_M(Y_i)-r^{(1)}(\varphi_M,t)\right)\Delta_i(t,D_n^+)-\mathbb{E}\left(\left(\varphi_M(Y_i)-r^{(1)}(\varphi_M,t)\right)\Delta_i(t,D_n^+)\right)\right),$$

and

$$J_2(x)=\sqrt{k}\left[\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)-r^{(1)}(\varphi,t)\widehat{r}_{n,1}^{(1)}(1,t,D_n^+)-\mathbb{E}\left(\widehat{r}_{n,2}^{(1)}(\varphi_M,t,D_n^+)-r^{(1)}(\varphi,t)\widehat{r}_{n,1}^{(1)}(1,t,D_n^+)\right)\right].$$

Thus, the claimed result now is

$$J_2(x)\to\mathcal{N}\left(0,\sigma^2(t)\right).$$

The asymptotic normality of $J_2(x)$ was proved in Lemma 8.8 by choosing the bandwidth parameter as $u=D_n^+$. For $J_3(x)$ and $J_4(x)$, we obtain by using Lemma 8.2 and the fact that $\mathbb{E}\left(\widehat{r}_{n,1}^{(1)}(1,t,u)\right)=1$ with $u=D_n^-$or D_n^+

$$\begin{aligned}|J_3(x)+J_4(x)| =\ & \sqrt{k}\left(\left|r^{(1)}(\varphi,t)\left[\left(\widehat{r}_{n,1}^{(1)}(1,t,D_n^+)-\mathbb{E}\left(\widehat{r}_{n,1}^{(1)}(1,t,D_n^+)\right)\right)\right]\right|\right)\\ &+\sqrt{k}\left(\left|r^{(1)}(\varphi,t)\left[|\,\mathbb{E}\left(\widehat{r}_{n,1}^{(1)}(1,t,H_{n,k}(t))\right)-\widehat{r}_{n,1}^{(1)}(1,t,H_{n,k}(t))\right]\right|\right)\end{aligned}$$

$$\leq \sqrt{k}\left|r^{(1)}(\varphi,t)\left(\widehat{r}^{(1)}_{n,1}(1,t,D_n^+)-1\right)\right| + \sqrt{k}\left|r^{(1)}(\varphi,t)\left(1-\widehat{r}^{(1)}_{n,1}(1,t,D_n^+)\right)\right|$$
$$\leq 2\sqrt{k}\left|r^{(1)}(\varphi,t)\left(\widehat{r}^{(1)}_{n,1}(1,t,D_n^+)-1\right)\right|.$$

Using Lemma 8.5, we get

$$\sqrt{k}\left|\widehat{r}^{(1)}_{n,1}(1,t,D_n^+)-1\right| \to 0 \text{ as } n\to\infty.$$

Consequently, we have

$$|J_3(x)+J_4(x)| \to 0.$$

Hence, the proof is complete.

Proof of Lemma 8.12:

For the proof of this Lemma, it suffices to use the result of [58]) in inequality (A.9), Tchebychev's inequality and Lemma 8.2. For $\varepsilon>0$, we readily infer

$$\mathbb{P}\left(\left|\widehat{r}^{(1)}_{n,1}(1,t,H_{n,k}(t))-1\right|>\delta\right)$$
$$\leq \mathbb{P}\left(\left|\widehat{r}^{(1)}_{n,1}(1,t,D_n^-)-\mathbb{E}\left(\widehat{r}^{(1)}_{n,1}(1,t,D_n^-)\right)\right|>\varepsilon\right) \leq \frac{\operatorname{Var}\left(\widehat{r}^{(1)}_{n,1}(1,t,D_n^-)\right)}{\varepsilon^2}.$$

Making use of the fact that

$$\operatorname{Var}\left(\widehat{r}^{(1)}_{n,1}(1,t,D_n^-)\right) = O\left(\frac{1}{n\phi\left(x,D_n^-\right)}\right).$$

We finally obtain

$$\widehat{r}^{(1)}_{n,1}(1,t,H_{n,k}(t))-1 \to 0, \text{ in probability as } n\to\infty$$

Hence, the proof is complete. ■

Proof of Lemma 8.13:

We have

$$\widehat{r}^{(1)}_{n,2}(\varphi,t,h_K(t))-\mathbb{E}\left(\widehat{r}^{(1)}_{n,2}(\varphi,t,h_K(t))\right)$$
$$= \frac{1}{n\mathbb{E}[\Delta_1(t,h_K(t))]}\sum_{i=1}^{n}\left\{\varphi(Y_i)\Delta_i(t;h_K(t))-\mathbb{E}\left[\varphi(Y_i)\Delta_i(t;h_K(t))\right]\right\}$$
$$= \frac{1}{n\mathbb{E}[\Delta_1(t,h_K(t))]}\sum_{i=1}^{n} Z_i(t,h_K(t)),$$

where

$$Z_i(t,h_K(t)) = \varphi(Y_i)\Delta_i(t;h_K(t))-\mathbb{E}\left[\varphi(Y_i)\Delta_i(t;h_K(t))\right].$$

Taking into account condition **(C.4.2)** and using Hölder's inequality, for $2<q<p$ such that $\frac{1}{p}+\frac{1}{q}=1$, we can write for all $i\neq j$:

$$\mathbb{E}\left[\varphi(Y_i)\Delta_i(t;h_K(t))\varphi(Y_j)\Delta_j(t;h_K(t))\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\Delta_i(t;h_K(t))\Delta_j(t;h_K(t))\mathbb{E}\left[\varphi(Y_i)\varphi(Y_j) \mid (X_i,X_j)\right]\right] \\
&\leq \mathbb{E}\left[\Delta_i(t;h_K(t))\Delta_j(t;h_K(t))\mathbb{E}^{1/p}\left[F^p(Y) \mid X\right]\mathbb{E}^{1/q}\left[F^q(Y) \mid X\right]\right] \\
&\leq C\mathbb{E}\left[\Delta_i(t;h_K(t))\Delta_j(t;h_K(t))\right] \\
&\leq C\sup_{i\neq j}\mathbb{P}\left\{\left(X_i,X_j\right)\in B\left(t,h_K(t)\right)\times B\left(t,h_K(t)\right)\right\} \\
&\leq C\Psi(h_K(t))f_2(t), \qquad \text{(A.36)}
\end{aligned}
$$

the later inequality is due to condition **(C.1.2)**. Next, we have

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{1}{n\mathbb{E}[\Delta_1(t,h_K(t))]}\sum_{i=1}^{n}Z_i(t,h_K(t))\right) \\
&= \frac{1}{n^2\mathbb{E}^2\left(\Delta_1(t,h_K(t))\right)}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}\left(\varphi(Y_i)\Delta_i(t;h_K(t)),\varphi(Y_j)\Delta_j(t;h_K(t))\right) \\
&= \frac{1}{n\mathbb{E}^2\left(\Delta_1(t,h_K(t))\right)}\mathrm{Var}\left(\varphi(Y_1)\Delta_1(t;h_K(t))\right) \\
&\quad + \frac{1}{n^2\mathbb{E}^2\left(\Delta_1(t,h_K(t))\right)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\mathrm{Cov}\left(\varphi(Y_i)\Delta_i(t;h_K(t)),\varphi(Y_j)\Delta_j(t;h_K(t))\right) \\
&=: V_1+V_2.
\end{aligned}
$$

Let us start with the term V_1, considering the conditions **(C.1.1)** and **(C.3.1)**, it follows that

$$
\kappa_1^j C_1 f_1(t)\phi(h_K(t)) \leq \mathbb{E}\left(\Delta_1^j(t,h_K(t))\right) \leq \kappa_2^j C_2 f_1(t)\phi(h_K(t)), \qquad \text{(A.37)}
$$

and we have

$$
\begin{aligned}
\mathrm{Var}\left(\varphi(Y_1)\Delta_1(t;h_K(t))\right) &= \mathbb{E}\left[\varphi^2(Y_1)\Delta_1^2(t,h_K(t))\right]-\mathbb{E}^2\left[\varphi(Y_1)\Delta_1(t,h_K(t))\right] \\
&\leq \mathbb{E}\left[\varphi^2(Y_1)\Delta_1^2(t,h_K(t))\right] \\
&\leq \mathbb{E}\left[\mathbb{E}\left[F^2(Y)\mid X\right]\Delta_1^2(t,h_K(t))\right] \\
&\leq \mathbb{E}^{2/p}\left[F^p(Y)\mid X\right]\mathbb{E}^{2/q}\left[\Delta_1^q(t,h_K(t))\right] \\
&\leq C^{2/p}C_2^{2/q}\kappa_2^2\phi^2(h_K(t))f_1^2(t).
\end{aligned}
$$

Hence, combining the later inequality with (A.37) gives us:

$$
\frac{const1}{nf_1(t)\phi(h_K(t))} \leq V_1 \leq \frac{const2}{nf_1(t)\phi(h_K(t))}, \qquad \text{(A.38)}
$$

whenever $f_1(t)>0$ and $const1 < const2$. Next, we consider V_2. We have

$$
V_2 = \frac{1}{n^2\mathbb{E}^2\left(\Delta_1(t,h_K(t))\right)}\left\{\sum_{\substack{i=1\\ 0<|i-j|\leq\omega_n}}^{n}\sum_{j=1}^{n}\mathrm{Cov}\left(\varphi(Y_i)\Delta_i(t;h_K(t)),\varphi(Y_j)\Delta_j(t;h_K(t))\right)\right.
$$

$$
\begin{aligned}
& + \sum_{\substack{i=1 \\ |i-j|>\omega_n}}^{n} \sum_{j=1}^{n} \mathrm{Cov}\left(\varphi(Y_i)\Delta_i(t;h_K(t)), \varphi(Y_j)\Delta_j(t;h_K(t))\right)\Bigg\} \\
=: \quad & V_{2,1} + V_{2,2}, \qquad\qquad (A.39)
\end{aligned}
$$

whenever ω_n satisfies, $\omega_n = o(n)$, and by conditions **(C.1.2)–(C.3.1)**, we readily infer that:

$$
\mathrm{Cov}\left(\varphi(Y_i)\Delta_i(t;h_K(t)), \varphi(Y_j)\Delta_j(t;h_K(t))\right) \;\leq\; C\Psi(h_K(t))f_2(t) - C^{2/p}C_2^{2/q}\kappa_1^2\phi^2(h_K(t))f_1^2(t).
$$

Thus, by the previous equation in combination with (A.39), we have

$$
\begin{aligned}
V_{2,1} &\leq \frac{\Psi(h_K(t))f_2(t) - C^{2/p}C_2^{2/q}\kappa_1^2\phi^2(h_K(t))f_1^2(t)}{n^2\mathbb{E}^2\left(\Delta_1(t,h_K(t))\right)} n\omega_n \\
&= \frac{\kappa_2^2 f_2(t)\Psi(h_K)\omega_n}{n\mathbb{E}^2\left(\Delta_1(t,h_K(t))\right)} - C^{2/p}C_2^{2/q}\kappa_1^2\phi^2(h_K(t))f_1^2(t)\frac{\omega_n}{n}.
\end{aligned}
$$

Making use of (A.37), we obtain

$$
V_{2,1} \leq \frac{const' f_2(t)\Psi(h_K)\omega_n}{n f_1^2(t)\phi^2(h_K)} - const''\frac{\omega_n}{n}\phi^2(h_K(t))f_1^2(t).
$$

This when combined with (A.38) implies that

$$
\begin{aligned}
\frac{V_{2,1}}{V_1} &\leq const' \frac{f_2(t)}{f_1(t)}\frac{\Psi(h_K)\omega_n}{\phi(h_K)} - \omega_n\phi(h_K)f_1(t)C^{2/p}C_2^{2/q}\kappa_2^2\phi^2(h_K(t))f_1^2(t) \\
&\leq const' \frac{f_2(t)}{f_1(t)}\frac{\Psi(h_K)\omega_n}{\phi(h_K)} - const''\omega_n\phi^3(h_K)f_1^3(t),
\end{aligned}
$$

where ω_n is chosen in such a way that the above bound tends to 0 as $n \to \infty$. Now, let's consider $V_{2,2}$. For any two σ-algebras $\mathcal{A}$ and $\mathcal{B}$: $\alpha(\mathcal{A},\mathcal{B}) \subseteq \beta(\mathcal{A},\mathcal{B})$, so by applying Davydov's lemma on strong-mixing sequences, and taking into account condition **(C.4.2)**, we infer

$$
\begin{aligned}
& \mathrm{Cov}\left(\varphi(Y_i)\Delta_i(t;h_K(t)), \varphi(Y_j)\Delta_j(t;h_K(t))\right) \\
&\leq 8\left\{\mathbb{E}\left|\varphi(Y_i)\Delta_i(t;h_K(t))\right|^p\right\}^{2/p}\left[\beta(|i-j|)\right]^{1-2/p} \\
&\leq 8\left\{\mathbb{E}\left[\mathbb{E}\left[|F(Y_i)|^p \mid X\right]|\Delta_i(t;h_K(t))|^p\right]\right\}^{2/p}\left[\beta(|i-j|)\right]^{1-2/p} \\
&\leq 8C\left\{\mathbb{E}\left[|\Delta_i(t;h_K(t))|^p\right]\right\}^{2/p}\left[\beta(|i-j|)\right]^{1-2/p}.
\end{aligned}
$$

However, $\mathbb{E}\left(\Delta_i(t,h_K(t))\right)$ satisfies (A.37), then we have

$$
\mathrm{Cov}\left(\varphi(Y_i)\Delta_i(t;h_K(t)), \varphi(Y_j)\Delta_j(t;h_K(t))\right) \leq const f_1^{2/p}(t)\left\{\phi(h_K)\right\}^{2/p}\left[\beta(|i-j|)\right]^{1-2/p}.
$$

Which involves

$$
V_{2,2} \leq \frac{const f_1^{2/p}(t)\left\{\phi(h_K)\right\}^{2/p}}{n^2\mathbb{E}^2\left(\Delta_1(t,h_K(t))\right)} \sum_{\substack{i=1 \\ |i-j|>\omega_n}}^{n}\sum_{j=1}^{n}\left[\beta(|i-j|)\right]^{1-2/p}.
$$

Once more by using (A.37) and by a simple calculation (reduction of the double sum), we get:

$$V_{2,2} \leq \frac{const}{n\omega_n^{\delta} f_1^{2(1-1/p)}(t)\,\{\phi(h_K)\}^{2(1-1/p)}} \sum_{k=\omega_n+1}^{\infty} [k^{\delta}\beta(k)]^{1-2/p}.$$

This latter, with the boundary on V_1 implies:

$$\frac{V_{2,2}}{V_1} \leq \frac{const}{\omega_n^{\delta}\left(\log(w_n)\right)^{\delta(1-1/p)} f_1^{1-2/p}(t)\,\{\phi(h_K)\}^{1-2/p}} \sum_{k=\omega_n+1}^{\infty} k^{\delta}\left(\log(k)\right)^{\delta(1-1/p)} (\beta(k))^{1-2/p}. \tag{A.40}$$

Choosing $\omega_n = \{\phi(h_K)\}^{-(1-2/p)/\delta}$ and making use of condition **(C.5.1)**, we get that:

$$\begin{aligned}\frac{V_{2,2}}{V_1} \leq & \frac{const}{\left[-\frac{1}{\delta}(1-\frac{2}{p})\log(\phi(h_K))\right]^{\delta(1-1/p)} f_1^{1-2/p}(t)} \sum_{k=\omega_n+1}^{\infty} k^{\delta}\left(\log(k)\right)^{\delta(1-1/p)} (\beta(k))^{1-2/p} \\ & \underset{n\to\infty}{\longrightarrow} 0.\end{aligned} \tag{A.41}$$

By the same value chosen for ω_n:

$$\frac{V_{2,1}}{V_1} \leq \frac{const f_2(t)\Psi(h_K)}{f_1(t)\,\{\phi(h_K)\}^2}\omega_n\phi(h_K) + \{\phi(h_K)\}^{1-(1-2/p)/\delta},$$

by the fact that $\Psi(h_K)/\phi^2(h_K)$ is assumed to be bounded and $\phi(h_K)\omega_n \to 0$, then the first term of the last inequality tends to 0 as $n \to \infty$. Furthermore, the second term also tends to 0, since $1-(1-2/p)/\delta < 1$. Hence, the proof is complete.

Proof of Lemma 8.14:

Based on (4.8), we have

$$\mu_n(\varphi, \mathbf{t}, \mathbf{h}(\mathbf{t})) = \sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}\left\{mu_n^{(1)}\left(\pi_{1,m}G_{\varphi,\mathbf{t},\mathbf{h}}\right) + \sum_{p=2}^{m}\frac{m!}{(m-p)!}u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}\right)\right\}.$$

Remark that the first term of the last sum is an empirical process indexed by $m\mathcal{G}^{(1)}$, so from Section 4.1, we have

$$\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}mu_n^{(1)}\left(\pi_{1,m}G_{\varphi,\mathbf{t},\mathbf{h}}\right) \xrightarrow{d} \mathcal{N}\left(0, m^2\sigma_{\mathbf{t}}^2(\varphi,\varphi)\right). \tag{A.42}$$

So, it is enough to show, for $2 \leq p \leq m$, that

$$\sqrt{n\tilde{\phi}(\mathbf{h}(\mathbf{t}))}u_n^{(p)}\left(\pi_{p,m}G_{\varphi,\mathbf{t},\mathbf{h}}\right) \xrightarrow{\mathbb{P}} 0. \tag{A.43}$$

To avoid ambiguity, we consider the case where $m = 2$ (the other cases are treated in the same way). Then, we have

$$u_n^{(2)}\left(\pi_{2,m}G_{\varphi,\mathbf{t},\mathbf{h}}\right) = \frac{1}{n(n-1)}\sum_{1\leqslant<i<j\leqslant n}\pi_{2,m}G_{\varphi,\mathbf{t},\mathbf{h}}\left((X_i,Y_i),(X_j,Y_j)\right).$$

The projections being $\mathbb{P}$-canonical, then to show (A.43) it suffices to show that

$$\mathbb{E}\left[u_n^{(2)}\left(\pi_{2,m}G_{\varphi,\mathbf{t}}\right)\right]^2 \xrightarrow{\mathbb{P}} 0.$$

We have

$$\mathbb{E}\left[u_n^{(2)}\left(\pi_{2,m}G_{\varphi,\mathbf{t}}\right)\right]^2 = \frac{1}{n^2\,(n-1)^2}\sum_{1\leqslant<i_1<i_2\leqslant n}\;\sum_{1\leqslant<j_1<j_2\leqslant n}\mathbb{E}\left((i_1,i_2);(j_1,j_2)\right),$$

where

$$\mathbb{E}\left((i_1,i_2);(j_1,j_2)\right) = \mathbb{E}\left[\pi_{2,m}G_{\varphi,\mathbf{t}}\left((X_{i_1},Y_{i_1}),(X_{i_2},Y_{i_2})\right)\times\pi_{2,m}G_{\varphi,\mathbf{t}}\left((X_{j_1},Y_{j_2}),(X_{j_2},Y_{j_2})\right)\right].$$

We first remark that

$$\begin{aligned}
&\mathbb{E}\left[n^2(n-1)^2u_n^{(2)}\left(\pi_{2,m}G_{\varphi,\mathbf{t}}\right)\right]^2\\
&= \sum_{1\leq i_1<i_2\leq n}\;\sum_{1\leq j_1<j_2\leq n}\mathbb{E}((i_1,i_2),(j_1,j_2))\\
&\leq \left|\sum_{1\leq i_1<i_2<j_1<j_2\leq n}\mathbb{E}((i_1,i_2),(j_1,j_2))\right| + \left|\sum_{1\leq i_1<j_1<i_2<j_2\leq n}\mathbb{E}((i_1,i_2),(j_1,j_2))\right|\\
&\quad + \left|\sum_{1\leq i_1<j<j_2<i_2\leq n}\mathbb{E}((i_1,i_2),(j_1,j_2))\right| + \left|\sum_{1\leq i_1<i_2<j_2\leq n}\mathbb{E}((i_1,i_2),(i_2,j_2))\right|\\
&\quad + \left|\sum_{i_1,i_2=1}^{n}\mathbb{E}((i_1,i_2),(i_1,i_2))\right|.
\end{aligned}$$

Following similar reasoning to obtain (A.41), the following bound suffices to our need

$$\begin{aligned}
&\left|\sum_{1\leq i_1<i_2<j_1<j_2\leq n}\mathbb{E}((i_1,i_2),(j_1,j_2))\right|\\
&\leq \left|\sum_{\substack{1\leq i_1<i_2<j_1<j_2\leq n\\ i_2-i_1\geq j_2-j_1}}\mathbb{E}((i_1,i_2),(j_1,j_2))\right| + \left|\sum_{\substack{1\leq i_1<i_2<j_1<j_2\leq n\\ j_2-j_1>i_2-i_1}}\mathbb{E}((i_1,i_2),(j_1,j_2))\right|\\
&\leq const\frac{n^2}{\tilde{\phi}(\mathbf{h})^4}\sum_{s=1}^{\infty}s^{\delta}\left(\log(s)\right)^{\delta(1-1/p)}\left(\beta(s)\right)^{1-2/p}.
\end{aligned}$$

Treating the other terms as in Lemma 2 of [203] (refer for detailed proof in the conditional setting to [13, Lemma 3]) we get

$$\mathbb{E}\left[u_n^{(2)}\left(\pi_{2,m}G_{\varphi,\mathbf{t}}\right)\right]^2 = O\left(n\tilde{\phi}(\mathbf{h})\right)^{-2}\longrightarrow 0.$$

Hence, the proof of (8.88) is complete. The statement (8.89) is a direct consequence of (8.88) in connection with condition **(C.8)**.

Proof of Lemma 8.15

For any two constants the sum of the product of these constants by the components of the vector (8.90) is a centered U-statistic, i.e., for $C^{(1)}, C^{(2)} \in \mathbb{R}$,

$$C^{(1)}\left[u_n(\varphi_1, \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(\varphi_1, \mathbf{t}, \mathbf{h})\right)\right] + C^{(2)}\left[u_n(\varphi_2, \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(\varphi_2, \mathbf{t}, \mathbf{h})\right)\right] = u_n(\varphi', \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(\varphi', \mathbf{t}, \mathbf{h})\right),$$

where

$$\varphi' = C^{(1)}\varphi_1 + C^{(2)}\varphi_2.$$

So, by the Cramér-Wold theorem the proof of the present lemma is directly deduced from Lemma 8.14. What remains to be done is to complete the proof of the theorem in question based on the two previous lemmas. Indeed,

$$\widehat{r}_n^{(m)}(\varphi, \mathbf{t}; \mathbf{h}) - \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right) = \frac{u_n(\varphi, \mathbf{t}, \mathbf{h}) - u_n(1, \mathbf{t}, \mathbf{h})\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)}{u_n(1, \mathbf{t}, \mathbf{h})},$$

and because

$$u_n(\varphi, \mathbf{t}, \mathbf{h}) \xrightarrow{\mathbb{P}} 1,$$

then all we have to do is prove that

$$\sqrt{n\tilde{\phi}(\mathbf{h})}\left\{u_n(\varphi, \mathbf{t}, \mathbf{h}) - u_n(1, \mathbf{t}, \mathbf{h})\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right\} \xrightarrow{d} \mathcal{N}\left(0, \rho^2\right). \tag{A.44}$$

We have

$$\begin{aligned} &\sqrt{n\tilde{\phi}(\mathbf{h})}\left\{u_n(\varphi, \mathbf{t}, \mathbf{h}) - u_n(1, \mathbf{t}, \mathbf{h})\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right\} \\ = \; &\left(1, \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right)\begin{pmatrix} \sqrt{n\tilde{\phi}(\mathbf{h})}\left\{u_n(\varphi, \mathbf{t}, \mathbf{h}) - \mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right)\right\} \\ \sqrt{n\tilde{\phi}(\mathbf{h})}\left\{u_n(1, \mathbf{t}, \mathbf{h}) - 1\right\} \end{pmatrix}, \end{aligned}$$

the latter, the fact that $r^{(m)}(\varphi, \mathbf{t})$ is continuous on $\mathbf{t}$ and $\mathbb{E}\left(u_n(\varphi, \mathbf{t}, \mathbf{h})\right) \to r^{(m)}(\varphi, \mathbf{t})$ leads to the desired result.

Proof of Lemma 8.16:

For the clarity of the exposition, we present the proof for $m = 2$; this case already contains the main idea. As in the proof of Theorem 4.2, we divide the sequence $\{(\mathbf{X}_i, \mathbf{Y}_i)\}$ into υ_n alternate blocks, here the sizes a_n, b_n are different satisfying

$$b_n \ll a_n, \quad (\upsilon_n - 1)(a_n + b_n) < n \leqslant \upsilon_n(a_n + b_n), \tag{A.45}$$

and set, for $1 \leqslant j \leqslant \upsilon_n - 1$:

$$\begin{aligned} H_j^{(U)} &= \left\{i : (j-1)(a_n + b_n) + 1 \leqslant i \leqslant (j-1)(a_n + b_n) + a_n\right\}, \\ T_j^{(U)} &= \left\{i : (j-1)(a_n + b_n) + a_n + 1 \leqslant i \leqslant (j-1)(a_n + b_n) + a_n + b_n\right\}, \\ H_{\upsilon_n}^{(U)} &= \left\{i : (\upsilon_n - 1)(a_n + b_n) + 1 \leqslant i \leqslant n \wedge (\upsilon_n - 1)(a_n + b_n) + a_n\right\}, \end{aligned}$$

$$T_{\upsilon_n}^{(U)} \quad = \quad \{i : (\upsilon_n - 1)(a_n + b_n) + a_n + 1 \leqslant i \leqslant n\}.$$

Note that the notation b_n used here and in the proof of Theorem 4.7 denotes the size of the alternative blocks. However, in the proof of Theorem 4.2, it denotes the radius of the nets of the class of functions. Then, we have

$$\begin{aligned}
&\sum_{i\neq j}^{n} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right) \\
&= \sum_{p\neq q}^{\upsilon_n} \sum_{i\in H_p^{(U)}} \sum_{j\in H_q^{(U)}} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right) \\
&\quad + \sum_{p=1}^{\upsilon_n} \sum_{i\neq j\, i,j\in H_p^{(U)}} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right) \\
&\quad + 2\sum_{p=1}^{\upsilon_n} \sum_{i\in H_p^{(U)}} \sum_{q:|q-p|\geqslant 2}^{\upsilon_n} \sum_{j\in T_q^{(U)}} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right) \\
&\quad + 2\sum_{p=1}^{\upsilon_n} \sum_{i\in H_p^{(U)}} \sum_{q:|q-p|\leqslant 1}^{\upsilon_n} \sum_{j\in T_q^{(U)}} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right) \\
&\quad + \sum_{p\neq q}^{\upsilon_n} \sum_{i\in T_p^{(U)}} \sum_{j\in T_q^{(U)}} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right) \\
&\quad + \sum_{p=1}^{\upsilon_n} \sum_{i\neq j\, i,j\in T_p^{(U)}} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right) \\
&=: \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI}. \qquad \text{(A.46)}
\end{aligned}$$

We have to treat each of the terms I–VI. The treatment of V and VI is readily achieved through similar techniques used to investigate I and II, which we omit.

The same type of block but not the same block (I):

Suppose that the sequence of independent blocks $\{\xi_i = (\varsigma_i, \zeta_i)\}_{i\in\mathbb{N}^*}$ is of size a_n. An application of (A.1), shows that

$$\begin{aligned}
&\mathbb{P}\left\{\left\|n^{-1/2}k^{-1} \sum_{p\neq q}^{\upsilon_n} \sum_{i\in H_p^{(U)}} \sum_{j\in H_q^{(U)}} \varphi(Y_i, Y_j) K\left(\frac{d\,(X_i, t_1)}{H_{n,k}(t_1)}\right) K\left(\frac{d\left(X_j, t_2\right)}{H_{n,k}(t_2)}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2} > \delta\right\} \\
&\leqslant \mathbb{P}\left\{\left\|n^{-1/2}k^{-1} \sum_{p\neq q}^{\upsilon_n} \sum_{i\in H_p^{(U)}} \sum_{j\in H_q^{(U)}} \varphi(\zeta_i, \zeta_j) K\left(\frac{d\,(\varsigma_i, t_1)}{H_{n,k}(t_1, \varsigma)}\right) K\left(\frac{d\left(\varsigma_j, t_2\right)}{H_{n,k}(t_2, \varsigma)}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2} > \delta\right\} \\
&\quad + 2\upsilon_n\beta_{b_n}.
\end{aligned}$$

We keep the choice of b_n and υ_n such that :

$$\upsilon_n b_n^r \leqslant 1, \tag{A.47}$$

which implies that $2\upsilon_n \beta_{b_n} \to 0$ as $n \to \infty$, so the term to consider is the second summand. For some $(\xi_{n,1}, \xi_{n,2}) \in (0,1)^2$, we choose $\mathbf{D}_n^- = (D_{n,1}^-, D_{n,2}^-)$ and $\mathbf{D}_n^+ = (D_{n,1}^+, D_{n,2}^+)$ such that for $j = 1, 2$

$$\phi\left(D_{n,j}^-\right) = \sqrt{\xi_{n,j}}\frac{k}{n}, \quad \phi\left(D_{n,j}^+\right) = \frac{k}{\sqrt{\xi_{n,j}}n} \quad \text{and} \quad \xi_n := \xi_{n,1}\xi_{n,2} \to 1,$$

then using the results of [44], we have $\mathbb{1}_{\left\{D_{n,j}^- \leq H_{n,k}(t_j) \leq D_{n,j}^+\right\}} \xrightarrow{a.co} 1$ when $\frac{k}{n} \to 0$ and making use of condition **(C.3.2)**, we have :

$$u_n(\varphi, \mathbf{t}; \mathbf{D}_n^+) \leq u_n(\varphi, \mathbf{t}; \mathbf{h}_{n,k}(t)) \leq u_n(\varphi, \mathbf{t}; \mathbf{D}_n^-), \tag{A.48}$$

where $\mathbf{D}_n^- = (D_{n,1}^-, D_{n,2}^-)$ and $\mathbf{D}_n^+ = (D_{n,1}^+, D_{n,2}^+)$. This implies that

$$\begin{aligned}
&\mathbb{P}\left\{\left\|n^{-1/2}k^{-1}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d\left(\varsigma_i,t_1\right)}{H_{n,k}(t_1,\varsigma)}\right)K\left(\frac{d\left(\varsigma_j,t_2\right)}{H_{n,k}(t_2,\varsigma)}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2} > \delta\right\} \\
&\leq \mathbb{P}\left\{\left\|n^{-1/2}k^{-1}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d\left(\varsigma_i,t_1\right)}{D_{n,1}^-}\right)K\left(\frac{d\left(\varsigma_j,t_2\right)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2} > \delta\right\}.
\end{aligned}$$

Now, combining Lemma A.1 of [67] with Proposition B.8 in the Appendix, we obtain

$$\begin{aligned}
&\mathbb{E}\left\|n^{-1/2}k^{-1}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d\left(\varsigma_i,t_1\right)}{D_{n,1}^-}\right)K\left(\frac{d\left(\varsigma_j,t_2\right)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2} \\
&\leqslant \quad c_2\mathbb{E}\left\|n^{-1/2}k^{-1}\sum_{p\neq q}^{\upsilon_n}\epsilon_p\epsilon_q\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d\left(\varsigma_i,t_1\right)}{D_{n,1}^-}\right)K\left(\frac{d\left(\varsigma_j,t_2\right)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2} \\
&\leqslant \quad c_2\mathbb{E}\int_0^{D_{n\mathbf{h}}^{(U_1)}} N\left(u, \mathscr{F}_2\mathscr{K}^2, \widetilde{d}_{nh_K,2}^{(1)}\right)du,
\end{aligned} \tag{A.49}$$

where $D_{n\mathbf{h}}^{(U_1)}$ is the diameter of $\mathscr{F}_2\mathscr{K}^2$ according to the distance $\widetilde{d}_{nh_K,2}^{(1)}$, which are defined respectively by

$$D_{n\mathbf{h}}^{(U_1)} := \left\|\mathbb{E}_\epsilon\left|n^{-1/2}k^{-1}\sum_{p\neq q}^{\upsilon_n}\epsilon_p\epsilon_q\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d\left(\varsigma_i,t_1\right)}{D_{n,1}^-}\right)K\left(\frac{d\left(\varsigma_j,t_2\right)}{D_{n,2}^-}\right)\right|\right\|_{\mathscr{F}_2\mathscr{K}^2},$$

and

$$\widetilde{d}_{nh_K,2}^{(1)}\left(\varphi_1\tilde{K}_1, \varphi_2\widetilde{K}_2\right)$$

$$:= \mathbb{E}_\epsilon \left| n^{-1/2}k^{-1}\sum_{p\neq q}^{\upsilon_n}\epsilon_p\epsilon_q\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\left[\varphi_1(\zeta_i,\zeta_j K_1\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K_1\left(\frac{d\left(\varsigma_j,t_2\right)}{D^-_{n,2}}\right)\right.\right.$$
$$\left.\left. - \varphi_2(\zeta_i,\zeta_j)K_2\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K_2\left(\frac{d\left(\varsigma_j,t_2\right)}{D^-_{n,2}}\right)\right]\right|.$$

Let consider another semi-norm $\widetilde{d}^{(2)}_{nh_K,2}$

$$\widetilde{d}^{(2)}_{nh,2}\left(\varphi_1\tilde{K}_1,\varphi_2\tilde{K}_2\right) = \frac{1}{n\tilde{\phi}(\mathbf{D}^-_n)}\left[\sum_{i\neq j}^{\upsilon_n}\left(\varphi_1(\zeta_i,\zeta_j)K_1\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K_1\left(\frac{d\left(\varsigma_j,t_2\right)}{D^-_{n,2}}\right)\right.\right.$$
$$\left.\left.-\varphi_2(\zeta_i,\zeta_j)K_2\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K_2\left(\frac{d\left(\varsigma_j,t_2\right)}{D^-_{n,2}}\right)\right)^2\right]^{1/2}.$$

One can see that we have

$$\widetilde{d}^{(1)}_{nh_K,2}\left(\varphi_1K_1,\varphi_2K_2\right) \leq a_n n^{1/2}k^{-3}\widetilde{d}^{(2)}_{nh,2}\left(\varphi_1K_1,\varphi_2K_2\right).$$

We readily infer that

$$\mathbb{E}\left\| n^{-1/2}k^{-1}\sum_{p\neq q}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{H_{n,k}(t_1,\varsigma)}\right)K\left(\frac{d\left(\varsigma_j,t_2\right)}{H_{n,k}(t_2,\varsigma)}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}$$
$$\leqslant c_2\mathbb{E}\int_0^{D^{(U_1)}_{nh_K}} N\left(ua_n^{-1}n^{-1/2}k^3,\mathscr{F}_2\mathscr{K}^2,\widetilde{d}^{(2)}_{nh_K,2}\right)du$$
$$\leqslant c_2a_nn^{1/2}k^{-3}\mathbb{P}\left\{D^{(U_1)}_{nh_K}a_n^{-1}n^{1/2}k^{-3}\geqslant\lambda_n\right\}+c_2a_nn^{-3/2}k^{-1}\int_0^{\lambda_n}\log u^{-1}du,$$

where $\lambda_n\to 0$. Notice that as $\lambda\to 0$, we have

$$\left(\int_0^{\lambda}\log u^{-1}du\right)\left(\lambda\log\lambda^{-1}\right)^{-1}\to 0,$$

where a_n and λ_n are chosen in such a way that the following relation will be fulfilled

$$a_n\lambda_nn^{1/2}k^{-3}\log\lambda_n^{-1}\to 0. \tag{A.50}$$

Making use of the triangle inequality, in combination with Hoeffding's trick, for instance, see [8, page 62], we obtain readily that

$$a_nn^{1/2}k^{-3}\mathbb{P}\left\{D^{(U_1)}_{nh_K}\geqslant\lambda_na_nn^{1/2}k^{-3}\right\}$$
$$\leqslant \lambda_n^{-2}a_n^{-1}n^{-3/2}k\mathbb{E}\left\|\sum_{p\neq q}^{\upsilon_n}\left[\sum_{i\in H_p^{(U)}}\sum_{j\in H_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K\left(\frac{d\left(\varsigma'_j,t_2\right)}{D^-_{n,2}}\right)\right]^2\right\|_{\mathscr{F}_2\mathscr{K}^2}$$

$$\leqslant \quad c_2 \upsilon_n \lambda_n^{-2} a_n^{-1} n^{-1/2} k \mathbb{E} \left\| \sum_{p=1}^{\upsilon_n} \left[\sum_{i,j \in H_p^{(U)}} \varphi(\zeta_i, \zeta_j') K\left(\frac{d(\varsigma_i, t_1)}{D_{n,1}^-} \right) K\left(\frac{d(\varsigma_j', t_2)}{D_{n,2}^-} \right) \right]^2 \right\|_{\mathscr{F}_2 \mathscr{K}^2}, \tag{A.51}$$

where $\left\{\xi_i' = (\varsigma_i', \zeta_i')\right\}_{i \in \mathbb{N}^*}$ are independent copies of $\{\xi_i = (\varsigma_i, \zeta_i)\}_{i \in \mathbb{N}^*}$. By imposing

$$\lambda_n^{-2} a_n^{1-r} n^{-1/2} k \to 0, \tag{A.52}$$

we readily infer that

$$\begin{aligned} &\left\| \upsilon_n \lambda_n^{-2} a_n^{-1} n^{-1/2} k \mathbb{E} \sum_{p=1}^{\upsilon_n} \left[\sum_{i,j \in H_p^{(U)}} \varphi(\zeta_i, \zeta_j') K\left(\frac{d(\varsigma_i, t_1)}{D_{n,1}^-} \right) K\left(\frac{d\left(\varsigma'_j, t_2\right)}{D_{n,2}^-} \right) \right]^2 \right\|_{\mathscr{F}_2 \mathscr{K}^2} \\ \leqslant \quad & O\left(\lambda_n^{-2} a_n^{1-r} n^{-1/2} k \right). \end{aligned}$$

By symmetrizing the expression in the expression in (A.51) and applying again the Proposition B.8 in the Appendix, we get

$$\begin{aligned} &\upsilon_n \lambda_n^{-2} a_n^{-1} n^{-1/2} k \mathbb{E} \left\| \sum_{p=1}^{\upsilon_n} \left[\sum_{i,j \in H_p^{(U)}} \epsilon_p \varphi(\zeta_i, \zeta_j') K\left(\frac{d(\varsigma_i, t_1)}{D_{n,1}^-} \right) K\left(\frac{d\left(\varsigma'_j, t_2\right)}{D_{n,2}^-} \right) \right]^2 \right\|_{\mathscr{F}_2 \mathscr{K}^2} \\ &\qquad \leqslant c_2 \mathbb{E} \left(\int_0^{D_{nh_K}^{(U_2)}} \left(\log N(u, \mathscr{F}_2 \mathscr{K}^2, \widetilde{d}'_{nh_K,2}) \right)^{1/2} \right), \end{aligned} \tag{A.53}$$

where

$$D_{nh_K}^{(U_2)} = \left\| \mathbb{E}_\epsilon \left| \upsilon_n \lambda_n^{-2} a_n^{-1} n^{-1/2} k \sum_{p=1}^{\upsilon_n} \epsilon_p \left[\sum_{i,j \in H_p^{(U)}} \varphi(\zeta_i, \zeta_j') K\left(\frac{d(\varsigma_i, t_1)}{D_{n,1}^-} \right) K\left(\frac{d\left(\varsigma'_j, t_2\right)}{D_{n,2}^-} \right) \right]^2 \right| \right\|_{\mathscr{F}_2 \mathscr{K}^2},$$

and for $\varphi_1 \widetilde{K}_1, \varphi_2 \tilde{K}_2 \in \mathscr{F}_2 \mathscr{K}^2$

$$\begin{aligned} &\widetilde{d}'_{nh_K,2}(\varphi_1 \tilde{K}_1, \varphi_2 \tilde{K}_2) \\ &\quad = \quad \mathbb{E}_\epsilon \left| \upsilon_n \lambda_n^{-2} a_n^{-1} n^{-1/2} k \sum_{p=1}^{\upsilon_n} \epsilon_p \left[\left(\sum_{i,j \in H_p^{(U)}} \varphi_1(\zeta_i, \zeta_j') K_1\left(\frac{d(\varsigma_i, t_1)}{D_{n,1}^-} \right) K_1\left(\frac{d(\varsigma_j', t_2)}{D_{n,2}^-} \right) \right)^2 \right. \right. \\ &\qquad \left. \left. - \left(\sum_{i,j \in H_p^{(U)}} \varphi_2(\zeta_i, \zeta_j') K_2\left(\frac{d(\varsigma_i, t_1)}{D_{n,1}^-} \right) K_2\left(\frac{d\left(\varsigma'_j, t_2\right)}{D_{n,2}^-} \right) \right)^2 \right] \right|. \end{aligned}$$

The fact that

$$\mathbb{E}_\epsilon \left| \upsilon_n \lambda_n^{-2} a_n^{-1} n^{-1/2} k \sum_{p=1}^{\upsilon_n} \epsilon_p \left(\sum_{i,j \in H_p^{(U)}} \varphi(\zeta_i, \zeta_j') K\left(\frac{d(\varsigma_i, t_1)}{D_{n,1}^-} \right) K\left(\frac{d\left(\varsigma'_j, t_2\right)}{D_{n,2}^-} \right) \right)^2 \right|$$

$$\leqslant a_n^{3/2}\lambda_n^{-2}n^{-1/2}\left[\upsilon_n^{-1}a_n^{-2}k^2\sum_{p=1}^{\upsilon_n}\sum_{i,j\in H_p^{(U)}}\left(\varphi(\zeta_i,\zeta'_j)K\left(\frac{d\left(\varsigma_i,t_1\right)}{D_{n,1}^-}\right)K\left(\frac{d\left(\varsigma'_j,t_2\right)}{D_{n,2}^-}\right)\right)^4\right]^{1/2},$$

so,

$$a_n^{3/2}\lambda_n^{-2}n^{-1/2}k^2\to 0, \tag{A.54}$$

we have the convergence of (A.53) to zero. For the choice of a_n, b_n, and υ_n, it should be noted that all the values satisfying (A.45), (A.47), (A.50), (A.52) and (A.54) are accepted.

The same blocks (II):

Remark that we have

$$\mathbb{P}\left\{\left\|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\varphi(Y_i,Y_j)K\left(\frac{d\left(X_i,t_1\right)}{H_{n,k}(t_1)}\right)K\left(\frac{d\left(X_j,t_2\right)}{H_{n,k}(t_2)}\right)\right\|_{\mathscr{F}\mathscr{K}^2}>\lambda\right\}$$
$$\leqslant\quad 2\upsilon_n\beta_{b_n}+\mathbb{P}\left\{\left\|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_2)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}>\lambda\right\}.$$

In a similar way as in the preceding proof, it suffices to prove that

$$\mathbb{E}\left(\left\|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,2}^-}\right)K\left(\frac{d(\varsigma_j,t_2)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\right)\to 0.$$

Because of computation by [138, p. 53] and the fact that the classes functions are uniformly bounded, we obtain uniformly in $\mathscr{F}_2\mathscr{K}^2$

$$\mathbb{E}\left(\sum_{i\neq j\, i,j\in H_p^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_2)}{D_{n,2}^-}\right)\right)=O(a_n).$$

This implies that we have to prove that

$$\mathbb{E}\quad\left(\left\|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\left[\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right.\right.\right.$$
$$\left.\left.\left.-\ \mathbb{E}\left(\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right)\right]\right\|_{\mathscr{F}_2\mathscr{K}^2}\right)\to 0. \tag{A.55}$$

Like for empirical processes, to prove (A.55), it suffices to symmetrize and show that

$$\mathbb{E}\left(\left\|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\epsilon_p\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\right)\to 0.$$

In a similar way as in (A.49), we infer that

$$\mathbb{E}\left(\left\|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\epsilon_p\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\right)$$
$$\leqslant \mathbb{E}\left(\int_0^{D_{nh_K}^{(U_3)}}\left(\log N\left(u,\mathscr{F}_2\mathscr{K}^2,\widetilde{d}_{nh_K,2}^{(3)}\right)\right)^{1/2}du\right),$$

where

$$D_{nh_K}^{(U_3)}=\left\|\mathbb{E}_\epsilon\left|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\epsilon_p\sum_{i\neq j\, i,j\in H_p^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right|\right\|_{\mathscr{F}_2\mathscr{K}^2},$$

and the semi-metric $\widetilde{d}_{nh_K,2}^{(3)}$ is defined by

$$\begin{aligned}&\widetilde{d}_{nh_K,2}^{(3)}\left(\varphi_1\widetilde{K}_1,\varphi_2\widetilde{K}_2\right)\\ &= \mathbb{E}_\epsilon\left|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\epsilon_p\sum_{i\neq j\, i,j\in H_p^{(U)}}\left(\varphi_1(\zeta_i,\zeta_j)K_1\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K_1\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right.\right.\\ &\quad\left.\left.-\ \varphi_2(\zeta_i,\zeta_j)K_2\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K_2\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right)\right|.\end{aligned}$$

Since we are trading uniformly bounded classes of functions, we infer that

$$\begin{aligned}&\mathbb{E}_\epsilon\left|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\epsilon_p\sum_{i\neq j\, i,j\in H_p^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right|\\ &\leqslant a_n^{3/2}n^{-1/2}k^{-1}\left[\frac{1}{\upsilon_n a_n^2}\sum_{p=1}^{\upsilon_n}\sum_{i\neq j\, i,j\in H_p^{(U)}}\left(\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right)^2\right]^{1/2}\\ &\leqslant O\left(a_n^{3/2}n^{-1/2}k^{-1}\right).\end{aligned}$$

Since $a_n^{3/2}(n)^{-1/2}k^{-1}\to 0$, $D_{nh_K}^{(U_3)}\to 0$, we obtain II $\to 0$ as $n\to\infty$.

Different types of blocks (III)

An application of (A.1), shows that

$$\begin{aligned}&\sum_{p=1}^{\upsilon_n}\mathbb{E}\left\|n^{-1/2}k^{-1}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(Y_i,Y_j)K\left(\frac{d(X_i,t_1)}{h_K}\right)K\left(\frac{d(X_j,t_2)}{h_K}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\\ &\leqslant \sum_{p=1}^{\upsilon_n}\mathbb{E}\left\|n^{-1/2}k^{-1}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\end{aligned}$$

$$+n^{-1/2}k^{-1}\upsilon_n^2 a_n b_n \beta_{a_n},$$

by the last choice of the parameters a_n, b_n, υ_n and the condition (8.92) imposed on the β-coefficients, we have

$$n^{-1/2}k^{-1}\upsilon_n^2 a_n b_n \beta_{a_n} \to 0.$$

For $p = 1$ and $p = \upsilon_n$, since we have independent exchangeable blocks, we infer that

$$\begin{aligned}
&\mathbb{E}\left\| n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\\
&= \mathbb{E}\left\| n^{-1/2}k^{-1}\sum_{i\in H_{\upsilon_n}^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\\
&= \mathbb{E}\left\| n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}.
\end{aligned}$$

For $2 \leqslant p \leqslant \upsilon_n - 1$, we obtain

$$\begin{aligned}
&\mathbb{E}\left\| n^{-1/2}k^{-1}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\geqslant 2}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\\
&= \mathbb{E}\left\| n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=4}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\\
&\leqslant \mathbb{E}\left\| n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2},
\end{aligned}$$

therefore, it suffices to treat the convergence

$$\mathbb{E}\left\| \upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2} \longrightarrow 0.$$

By similar arguments as in [8], the usual symmetrization gives

$$\begin{aligned}
&\mathbb{E}\left\| \upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\\
&\leqslant 2\mathbb{E}\left\| \upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\epsilon_q\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}
\end{aligned}$$

$$
\begin{aligned}
&= 2\mathbb{E}\left\{\left\|\upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\epsilon_q\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\mathbb{1}_{\left\{D_{nh_K}^{(U_4)}\leqslant\gamma_n\right\}}\right\}\\
&\quad+2\mathbb{E}\left\{\left\|\upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\epsilon_q\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\mathbb{1}_{\left\{D_{nh_K}^{(U_4)}>\gamma_n\right\}}\right\}\\
&= 2\mathrm{III}_1+2\mathrm{III}_2,
\end{aligned}
\tag{A.56}
$$

where

$$
D_{nh_K}^{(U_4)}=\left\|\upsilon_n n^{-1/2}k^{-1}\left[\sum_{q=3}^{\upsilon_n}\left(\sum_{j\in T_q^{(U)}}\sum_{i\in H_1^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right)^2\right]^{1/2}\right\|_{\mathscr{F}_2\mathscr{K}^2}.
$$

In a similar way as in (A.49), we infer that

$$
\mathrm{III}_1\leqslant c_2\int_0^{\gamma_n}\left(\log N\left(u,\mathscr{F}_2\mathscr{K}^2,\widetilde{d}_{nh_K,2}^{(4)}\right)\right)^{1/2}du, \tag{A.57}
$$

where

$$
\begin{aligned}
&\widetilde{d}_{nh_K,2}^{(4)}(\varphi_1\widetilde{K}_1,\varphi_2\widetilde{K}_2)\\
&\quad= \mathbb{E}_\epsilon\left|\upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\epsilon_q\left[\varphi_1(\zeta_i,\zeta_j)K_1\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K_1\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right.\right.\\
&\qquad\left.\left.-\ \varphi_2(\zeta_i,\zeta_j)K_2\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K_2\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right]\right|.
\end{aligned}
$$

Since we have

$$
\begin{aligned}
&\mathbb{E}_\epsilon\left|\upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\epsilon_q\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right|\\
&\quad\leqslant a_n^{-1/2}b_n n^{1/2}k\left[\left(\frac{1}{a_nb_n\upsilon_n n^{-4}k^4}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\left[\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right]^2\right)^{1/2}\right],
\end{aligned}
$$

and by considering the semi-metric

$$
\begin{aligned}
&\widetilde{d}_{nh_K,2}^{(5)}(\varphi_1\widetilde{K}_1,\varphi_2\widetilde{K}_2)\\
&\quad= \left(\frac{1}{a_nb_n\upsilon_n\tilde{\phi}^4(\mathbf{D}_n^-)}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\left[\varphi_1(\zeta_i,\zeta_j)K_1\left(\frac{d(\varsigma_i,t_1)}{D_{n,1}^-}\right)K_1\left(\frac{d(\varsigma_j,t_1)}{D_{n,2}^-}\right)\right.\right.
\end{aligned}
$$

$$
-\ \varphi_2(\zeta_i,\zeta_j)K_2\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K_2\left(\frac{d(\varsigma_j,t_1)}{D^-_{n,2}}\right)\Bigg]^2\Bigg)^{1/2}.
$$

We show that the expression in (A.57) is bounded as follows

$$
\upsilon_n^{1/2}b_n n^{-1/2}\tilde{\phi}(\mathbf{D}_n^-)\int_0^{\upsilon_n^{-1/2}b_n^{-1}n^{1/2}h^{-1}\gamma_n}\left(\log N\left(u,\mathscr{F}_2\mathscr{K}^2,\widetilde{d}^{(5)}_{nh_K,2}\right)\right)^{1/2}du,
$$

by choosing $\gamma_n = n^{-\alpha}$ for some $\alpha > (17r-26)/60r$, we get the convergence to zero of the previous quantity. To bound the second term in the right-hand side of (A.56), we remark that

$$
\begin{aligned}
&\mathrm{III}_2 && \text{(A.58)}\\
&= \mathbb{E}\left\{\left\|\upsilon_n n^{-1/2}k^{-1}\sum_{i\in H_1^{(U)}}\sum_{q=3}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\epsilon_q\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K\left(\frac{d(\varsigma_j,t_1)}{D^-_{n,2}}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}\mathbb{1}_{\left\{D^{(U_4)}_{nh_K}>\gamma_n\right\}}\right\}\\
&\leqslant a_n^{-1}b_n n^{1/2}\tilde{\phi}^{-1}(\mathbf{D}_n^-)\mathbb{P}\left\{\left\|\upsilon_n^2 n^{-3}\tilde{\phi}^{-2}(\mathbf{D}_n^-)\sum_{q=3}^{\upsilon_n}\left(\sum_{j\in T_q^{(U)}}\sum_{i\in H_1^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K\left(\frac{d(\varsigma_j,t_1)}{D^-_{n,2}}\right)\right)^2\right\|_{\mathscr{F}_2\mathscr{K}^2}\right.\\
&\qquad\left.\geqslant\gamma_n^2\right\}. && \text{(A.59)}
\end{aligned}
$$

We now apply the square root trick to the last expression conditionally on H_1^U. We denote by $\mathbb{E}_T$ the expectation with respect to $\sigma\left\{\left(\varsigma_j,\zeta_j\right): j\in T_q, q\geqslant 3\right\}$ and we get by **(C.6.)**, for $2r/(r-1) < s < \infty$, (in the notation in Lemma 5.2 of [102])

$$
\begin{aligned}
M_n &= \upsilon_n^{1/2}\mathbb{E}_T\left(\sum_{j\in T_q^{(U)}}\sum_{i\in H_1^{(U)}}\varphi(\zeta_i,\zeta_j)K\left(\frac{d(\varsigma_i,t_1)}{D^-_{n,1}}\right)K\left(\frac{d(\varsigma_j,t_1)}{D^-_{n,2}}\right)\right)^2,\\
t &= \gamma_n^2 a_n^{5/2}n^{1/2}\tilde{\phi}^{-1}(\mathbf{D}_n^-),\ \ \rho=\lambda=2^{-4}\gamma_n a_n^{5/4}n^{1/4}\tilde{\phi}^{-1/2}(\mathbf{D}_n^-)),\ \ m=\exp\left(\gamma_n^2 n\tilde{\phi}^{-2}(\mathbf{D}_n^-)b_n^{-2}\right).
\end{aligned}
$$

Since we need $t > 8M_n$, and $m\to\infty$, by similar arguments as in [8] page 69, we get the convergence of (A.57) and (A.59) to zero.

Different types of blocks (IV)

We have :

$$
\left\|n^{-1/2}k^{-1}\sum_{p=1}^{\upsilon_n}\sum_{i\in H_p^{(U)}}\sum_{q:|q-p|\leqslant 1}^{\upsilon_n}\sum_{j\in T_q^{(U)}}\varphi(Y_i,Y_j)K\left(\frac{d(X_i,t_1)}{H_{n,k}(t_1)}\right)K\left(\frac{d(X_j,t_2)}{H_{n,k}(t_2)}\right)\right\|_{\mathscr{F}_2\mathscr{K}^2}
\leqslant c_2\upsilon_n a_n b_n n^{-1/2}k^{-1}\to 0.
$$

Hence, the proof of the lemma is complete.

B. Auxiliary results

This appendix contains supplementary information that is an essential part of providing a more comprehensive understanding of the paper.

In the sequel, we define $X, X_1, \ldots, X_n$ to be i.i.d. random variables defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and taking values in some measurable space $(\Psi, \mathcal{B})$, and $\mathscr{F}$ to be a $\mathbb{P}$-measurable class of measurable functions with envelope function F, such that :

$$\mathbb{E}(F^2(X)) < \infty.$$

We further assume that $\mathscr{F}$ has the following property:

- For any sequence of i.i.d. $\mathcal{X}$-valued random variables $Z_1, Z_2, \ldots$ it holds that

$$\mathbb{E}\left\|\sum_{i=1}^{k}\{g(Z_i) - \mathbb{E}g(Z_1)\}\right\|_{\mathcal{G}} \leq C_1\sqrt{k}\,\|G(Z_1)\|_2, \quad 1 \leq k \leq n,$$

  where $C_1 > 1$ is a constant depending on $\mathcal{G}$ only.

Lemma B.1 (Theorem 2.14.1 [193]). *For an empirical process $\alpha_n(f)$ indexed by the class of functions $\mathscr{F}$ with the notation:*

$$\|\alpha_n(f)\|_{\mathscr{F}} := \sup_{f\in\mathscr{F}} |\alpha_n(f)|,$$

and $J(\delta, \mathscr{F})$ meaning

$$\sup_Q \int_0^\delta \sqrt{1 + \log N(\epsilon, \mathscr{F}, \|\cdot\|_{L_2(Q)})}d\epsilon,$$

we have, for $p > 1$,

$$\left\|\|\alpha_n(f)\|_{\mathscr{F}}\right\|_p \leq CJ(\delta, \mathscr{F})\,\|F\|_{p\vee 2}\,.$$

Lemma B.2 (Theorem 3.1 [78]). *Let $\mathscr{F}$ be a pointwise measurable function class satisfying the above assumptions. If we suppose that the empirical process $\alpha_n(f)$ satisfies:*

$$\mathbb{E}\,\|\alpha_n(f)\|_{\mathscr{F}} \leq C\|F\|_2, \tag{B.1}$$

then for any measurable subset $B \in \mathcal{B}$:

$$\mathbb{E}\,\|\alpha_n(f\mathbb{1}_B)\|_{\mathscr{F}} \leq 2C\|F\mathbb{1}_B\|_2.$$

From Theorem 3.2 of [78] it follows that a VC-type class of functions satisfies, always, the condition (B.1).

Lemma B.3 (Bernstein type inequality Fact 4.2 [78]). *Assume that for some $H > 0$ and $p \geq 2$ the r.vs $X, X_1, \ldots, X_n$ satisfy:*

$$\mathbb{E}(F^p(X)) \leq (p!/2)\sigma^2 H^{p-2}, \text{ where } \sigma^2 \geq \mathbb{E}(F^2(X)),$$

then for $\mu_n' = \mathbb{E}\left(\left\| \sqrt{n}\alpha_n(f)\right\|_{\mathscr{F}}\right)$, we have for any $z > 0$:

$$\mathbb{P}\left\{\max_{1\leq k\leq n}\left\| \sqrt{k}\alpha_k(f)\right\|_{\mathscr{F}} \geq \mu_n' + z\right\} \leq \exp\left(\frac{-z^2}{2n\sigma^2 + 2zH}\right)$$
$$\leq \exp\left(\frac{-z^2}{4n\sigma^2}\right) \vee \exp\left(\frac{-z}{4H}\right).$$

Lemma B.4 (Proposition 4. [9])**.** *Let $X_1, \ldots, X_m$ be i.i.d. random variables with values in some measurable space $(S, \mathcal{S})$. Let $\mathscr{F}$ be a class of symmetric functions f from S^m satisfying some measurability conditions. Suppose that there exists a finite constant c_0 such that for each $x_j \in S$ we have:*

$$\mathbb{E}\left\{\left\| n^{1-m}\sum_{I_n^m}\zeta_{i_1}\zeta_{i_2}f(x_{i_1}, \ldots, x_{i_m})\right\|_{\mathscr{F}}\right\} \leq c_0,$$

and that there is a finite constant b such that $\sup_{f\in\mathcal{F}}|f(X)| \leq b$, a.s. Then for each $u > 0$:

$$\mathbb{P}\left\{n^{1/2}\left\|\sum_{j=2}^{m}\binom{m}{j}u_n\left(\pi_{j,m}f\right)\right\|_{\mathscr{F}} \geq u\right\} \leq 2\exp\left(-\frac{un^{1/2}}{2^{m+5}m^{m+1}bc_0}\right),$$

where the variables ζ_i are a Rademacher variables.

Lemma B.5 (Lemma 6.1 in [124], p.186)**.** *Let $X_1, \ldots, X_m$ be independent Bernoulli random variables with $\mathbb{P}(X_i = 1) = p$, for all $i = 1, \ldots, n$. Set $U = X_1 + \cdots + X_m$ and $\mu = pn$. Then, for any $\omega \geq 1$, we have*

$$\mathbb{P}(U \geq (1+\omega)\mu) \leq \exp\left\{\frac{-\mu\min(\omega^2, \omega)}{4}\right\},$$

and if $\omega \in (0, 1)$, we have

$$\mathbb{P}\{U \leq (1-\omega)\mu\} \leq \exp\left\{-\mu\left(\omega^2/2\right)\right\}.$$

Lemma B.6. *Let $\{X_n\}_{n\in\mathbb{Z}}$ be a data sequence, along with the kernel function $h(\cdot)$, satisfying Assumptions (A1)–(A3). We then have, there exist absolute constants $C_4, C_5 > 0$ only depending on γ and r, such that, for any $x \geq 0$ and T sufficiently large,*

$$\mathbb{P}\left(\left|u_n^{(m)}(h) - \theta(h)\right| \geq C_4 M/\sqrt{n} + x\right) \leq 2\exp\left(-\frac{C_5 x^2 n}{M^2 + Mx(\log n)(\log\log 4n)}\right).$$

Proposition B.7. *Let $\{Z_{n,i}, 1 \leqslant i \leqslant n, n \geqslant 1\}$ a strong mixing non-stationary sequence of random variables, and let V be a $\sigma(Z_{n,i}, 1 \leqslant i \leqslant j)$-measurable and W be a $\sigma(Z_{n,i}, i \geqslant j+m)$-measurable. If $\mathbb{E}|V|^p < \infty$ and $\mathbb{E}|W|^q < \infty$, then:*

$$\text{Cov}(V, W) \leq [\alpha(m)]^{1/r} \{\mathbb{E}|V|^p\}^{1/p} \{\mathbb{E}|W|^q\}^{1/q},$$

where $p, q, r > 0$:

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1.$$

Proposition B.8 (Proposition 3.6 of [10])**.** *Let $\{\mathbf{X}_t : t \in \boldsymbol{T}\}$ be a process satisfying, for $m \geqslant 1$:*

$$(\mathbb{E}\|\mathbf{X}_t - \mathbf{X}_s\|^p)^{1/p} \leqslant \left(\frac{p-1}{q-1}\right)^{m/2} (\mathbb{E}\|\mathbf{X}_t - \mathbf{X}_s\|^q)^{1/q}, \quad 1 < q < p < \infty,$$

and the semi-metric : $\rho(s,t) = \left(\mathbb{E}\|\mathbf{X}_t - \mathbf{X}_s\|^2\right)^{1/2}$. There exists a constant $K = K(m)$ such that:

$$\mathbb{E}\sup_{s,t\in\mathbf{T}} \|\mathbf{X}_t - \mathbf{X}_s\| \leqslant K \int_0^D [\log N(\epsilon, \mathbf{T}, \rho)]^{m/2} d\epsilon,$$

where D being the ρ-diameter of $\mathbf{T}$.

Remark B.9. *In a similar way as in [35], Theorem 4.2 can be used to investigate the following problems.*

1) (Expectile regression). For $p \in (0,1)$, let $\psi(T - \boldsymbol{\theta}) = (p - \{T - \boldsymbol{\theta} \leq 0\})|T - \boldsymbol{\theta}|$, then the zero of $m(\psi, \cdot)$ with respect to $\boldsymbol{\theta}$ leads to quantities called expectiles by [155]. Expectiles, as defined by [155], may be introduced either as a generalization of the mean or as an alternative to quantiles. Indeed, classical regression provides us with a high sensitivity to extreme values, allowing for more reactive risk management. Quantile regression, on the other hand, provides the ability to acquire exhaustive information on the effect of the explanatory variable on the response variable by examining its conditional distribution, refer to [3, 4, 150, 151] for further details on expectiles in functional data setting.
2) (Quantile regression). For $p \in (0,1)$, let $\boldsymbol{\psi}(T - \boldsymbol{\theta}) = p - \{T - \boldsymbol{\theta} \leq 0\}$. Then the zero of $m(\psi, \cdot)$ with respect to $\boldsymbol{\theta}$ is the conditional p-th quantile, initially introduced in [129] in the real and linear framework, for more general setting, refer to [35].
3) (Conditional winsorized mean). As in [117], if we consider $\boldsymbol{\psi}(T - \boldsymbol{\theta}) = -k, T - \boldsymbol{\theta}, k$ if $T - \boldsymbol{\theta} < -k$, $|T - \boldsymbol{\theta}| \leq k$, or $T - \boldsymbol{\theta} > k$, then the zero of $m(\psi, \cdot)$ with respect to $\boldsymbol{\theta}$ will be the conditional winsorized mean. Notably, this parameter was not considered in the literature on nonparametric functional data analysis involving wavelet estimators. Our paper offers asymptotic results for the conditional winsorized mean when the covariates are functions.