



HAL
open science

Non-Parametric Conditional U-Processes for Locally Stationary Functional Random Fields under Stochastic Sampling Design

Salim Bouzebda, Inass Soukarieh

► **To cite this version:**

Salim Bouzebda, Inass Soukarieh. Non-Parametric Conditional U-Processes for Locally Stationary Functional Random Fields under Stochastic Sampling Design. *Mathematics*, 2023, 11 (1), pp.70. 10.3390/math11010016 . hal-03909074

HAL Id: hal-03909074

<https://hal.science/hal-03909074>

Submitted on 1 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Non-Parametric Conditional U -Processes for Locally Stationary Functional Random Fields under Stochastic Sampling Design

Salim Bouzebda ^{*,†}  and Inass Soukarieh [†] 

Laboratory of Applied Mathematics of Compiègne(LMAC), Université de Technologie de Compiègne, CEDEX, 60 203 Compiègne, France

* Correspondence: salim.bouzebda@utc.fr

† These authors contributed equally to this work.

Abstract: Stute presented the so-called conditional U -statistics generalizing the Nadaraya–Watson estimates of the regression function. Stute demonstrated their pointwise consistency and the asymptotic normality. In this paper, we extend the results to a more abstract setting. We develop an asymptotic theory of conditional U -statistics for locally stationary random fields $\{X_{s,A_n} : s \text{ in } R_n\}$ observed at irregularly spaced locations in $R_n = [0, A_n]^d$ as a subset of \mathbb{R}^d . We employ a stochastic sampling scheme that may create irregularly spaced sampling sites in a flexible manner and includes both pure and mixed increasing domain frameworks. We specifically examine the rate of the strong uniform convergence and the weak convergence of conditional U -processes when the explicative variable is functional. We examine the weak convergence where the class of functions is either bounded or unbounded and satisfies specific moment conditions. These results are achieved under somewhat general structural conditions pertaining to the classes of functions and the underlying models. The theoretical results developed in this paper are (or will be) essential building blocks for several future breakthroughs in functional data analysis.

Keywords: conditional U -statistics; locally stationary random field; functional data; empirical processes; conditional U -processes; VC-class of functions; kernel-type estimators; regression; irregularly spaced data

MSC: 60F05; 62G05; 62G08; 62G20; 62G35; 62G07; 62G32; 62G30; 62E20; 62G08; 62M10



Citation: Bouzebda, S.; Soukarieh, I. Non-Parametric Conditional U -Processes for Locally Stationary Functional Random Fields under Stochastic Sampling Design. *Mathematics* **2023**, *11*, 16. <https://doi.org/10.3390/math11010016>

Academic Editor: Elvira Di Nardo

Received: 21 October 2022

Revised: 14 December 2022

Accepted: 16 December 2022

Published: 20 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The regression problem has been studied by statisticians and probability theorists for many years, resulting in a vast array of approaches. Various themes have been covered, such as modeling, estimate method applications, tests, and other related topics. In addition to the parametric framework, in which one must estimate a finite number of parameters based on an a priori specified model structure, the non-parametric framework is devoted to data that lack a priori structural information. As inherent disadvantages, non-parametric processes are susceptible to estimation biases and reductions in convergence rates compared to parametric methods. Kernel non-parametric function estimation techniques have long been of great interest; for good references to research literature and statistical applications in this area, see [1–6] and the references therein. Even though they are widely used, they are just one of several possible approaches to building reliable function estimators. Despite their popularity, methods such as nearest neighbor, spline, neural network, and wavelet analysis are examples of these approaches. These techniques have been utilized on a vast range of different types of data. In this article, our focus will be narrowed to the development of consistent kernel-type estimators for the conditional U -statistics in the context of spatial data. Spatial data are typically generated in numerous research fields, such as econometrics, epidemiology, environmental science, image analysis, oceanography,

meteorology, geostatistics, etc. These data are typically collected in various fields and treated statistically on measurement sites. Consult [7–10] as well as the references contained in these works to find reliable sources of references to the research literature in this area and discover some statistical applications. In the context of non-parametric estimation for spatial data, the existing papers are mostly concerned with estimating probability density and regression functions. Hence, we will cite some important references [11–15] and the references in which they are included. By considering the conditional U -processes, we give a more generic and abstract context based on this research. With many possible applications, the idea of U -statistics (introduced in a landmark work by [16]) and U -processes have attracted a great deal of interest over the past few decades. U -processes are effective for resolving intricate statistical issues: density estimation, non-parametric regression tests, and goodness-of-fit tests are among the examples. Specifically, U -processes emerge in statistics in a variety of contexts, such as the terms of higher order in von Mises expansions. In particular, U -statistics assist in the analysis of estimators and function estimators, with varying degrees of smoothness. For example, Ref. [17] aimed to analyze the product limit estimator for shortened data, so he employs almost sure uniform bounds for \mathbb{P} -canonical U -processes. In addition, Ref. [18] introduced two novel normality tests based on U -processes. Likewise, new tests for normality that use as test statistics weighted L_1 -distances between the standard normal density and local U -statistics based on standardized data were introduced by [19–21]. In addition, Ref. [22] challenged the estimate of the mean of multivariate functions under the assumption of possibly heavy-tailed distributions and presented the median-of-means based on U -statistics. The applications of U -processes in various statistical applications may also include tests for functions' qualitative features in non-parametric statistics (c.f. [23–25]), cross-validation for density estimation [26], and establishing the limiting distributions of M -estimators (see, e.g., Refs. [27–29]). Historically, Ref. [27] furnishes the necessary and sufficient criteria for the law of large numbers and the sufficient conditions for the central limit theorem for U -processes, equipped by [16,30,31], who provided (amongst others) the first asymptotic results for the case that the underlying random variables are independent and identically distributed. However, under weak dependency assumptions, asymptotic outcomes are illustrated in [32–34] or just lately in [35] and in a more general setting in [36–41]. The applicability of U -statistics in estimation and machine learning applications is comprehensive. We refer to the U -statistics with random kernels of divergent orders to [40,42–45]. Infinite-order U -statistics are helpful tools for creating simultaneous prediction intervals. These constructed intervals are important to quantify ensemble methods' uncertainty such as subbagging and random forests. For additional information on the topic, c.f. [46]. The MeanNN method estimation for differential entropy, which was first described by [47], is a remarkable instance of the U -statistic. A novel test statistic for goodness-of-fit tests was proposed by [48] using U -statistics. Using U -statistics, the conference [49] proposed a measure to quantify the level of clustering quality exhibited by a partition. The interested reader may refer to [50–52] for outstanding resources of references on the U -statistics. The book of [29] provides a profound and in-depth view of the notion of U -processes.

In this work, our primary focus is on the scenario, including spatial–functional data. We give an excerpt from [53]: “*Functional data analysis (FDA) is a branch of statistics concerned with the analysis of infinite-dimensional variables such as curves, sets, and images. It has undergone phenomenal growth over the past 20 years, stimulated in part by major advances in data collection technology that have brought about the “Big Data” revolution. Often perceived as a somewhat arcane area of research at the turn of the century, FDA is now one of the most active and relevant fields of investigation in data science.*” The reader is directed to the works of reference [54,55] for an overview of this subject area. These references include fundamental approaches to functional data analysis and a wide range of case studies from diverse disciplines, such as criminology, economics, archaeology, and neurophysiology. It is important to note that the extension of probability theory to random variables taking values in normed vector

spaces (for example, Banach and Hilbert spaces), including extensions of certain classical asymptotic limit theorems, predates the recent literature on functional data; the reader is referred to the book [56] for more information on this topic. Considering density and mode estimates for data with values in a normed vector space is the focus of the work presented by [57]. The problem of the curse of dimensionality, which occurs when functional data have too many dimensions, is discussed in this study, along with potential solutions to the issue. According to [55], non-parametric models were deemed useful in regression estimation. We could also refer to [58–60].

Modern empirical process theory has recently been applied to processing functional data. Ref. [61] provided the consistency rates of several conditional models, such as the regression function, the conditional cumulative distribution, the conditional density, and others, uniformly over a subset of the explanatory variable. Ref. [62] extended [63]’s UIB consistency to the ergodic setting. Ref. [64] considered the problem of local linear estimation of the regression function when the regressor is functional and showed strong convergence, with specified rates, uniformly in bandwidth parameters. Ref. [65] examined the k -nearest neighbors (kNN) estimate of the non-parametric regression model for strong mixing functional time series data and determined the uniform, almost complete convergence rate of the k NN estimator under some mild conditions. Ref. [66] treated the ergodic data and offered a variety of results related to the limiting distribution for the conditional mode in the functional setting; for recent references, c.f [38,67–72].

Ref. [73] raised a class of estimators for $r^{(m)}(\varphi, t)$, known as conditional U -statistics, attempted to generalize the Nadaraya–Watson regression function estimations. Foremost, we present Stute’s estimators. Consider the 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 $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$. Let $\varphi : \mathcal{Y}^m \rightarrow \mathbb{R}$ be a measurable function. In this study, the estimation of the conditional expectation, or regression function, is our primary concern:

$$r^{(m)}(\varphi, t) = \mathbb{E}(\varphi(Y_1, \dots, Y_m) \mid (\mathbf{X}_1, \dots, \mathbf{X}_m) = t), \text{ for } t \in \mathbb{R}^{dm}, \tag{1}$$

whenever it exists, i.e.,

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

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

$$\sup_{x \in \mathbb{R}^d} |K(x)| =: \kappa < \infty \text{ and } \int K(x) dx = 1. \tag{2}$$

Hence, the class of estimators for $r^{(m)}(\varphi, t)$, given by [73], is defined, for each $t \in \mathbb{R}^{dm}$, as follows:

$$\hat{r}_n^{(m)}(\varphi, t; h_n) = \frac{\sum_{i \in I_n^m} \varphi(Y_{i_1}, \dots, Y_{i_m}) K\left(\frac{t_1 - X_{i_1}}{h_n}\right) \dots K\left(\frac{t_m - X_{i_m}}{h_n}\right)}{\sum_{i \in I_n^m} K\left(\frac{t_1 - X_{i_1}}{h_n}\right) \dots K\left(\frac{t_m - X_{i_m}}{h_n}\right)}, \tag{3}$$

where

$$I_n^m = \{i = (i_1, \dots, i_m) : 1 \leq i_j \leq n \text{ and } i_j \neq i_r \text{ if } j \neq r\},$$

denotes the set of all m -tuples of different integers i_j between 1 and n and $\{h_n := h_n\}_{n \geq 1}$ is a sequence of positive constants that converge to zero with rate $nh_n^m \rightarrow \infty$.

For $m = 1$, the $r^{(m)}(\varphi, t)$ becomes

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

and the estimate of Stute will be transformed to the Nadaraya–Watson estimator of $r^{(1)}(\varphi, t)$.

Behind, Ref. [74] aimed to estimate the rate of uniform convergence in t of $\hat{r}_n^{(m)}(\varphi, t; h_n)$ to $r^{(m)}(\varphi, t)$. Meanwhile, the study of [75] developed the limit distributions of $\hat{r}_n^{(m)}(\varphi, t; h_n)$, by discussing and contrasting the findings of Stute. Correspondingly, under appropriate mixing conditions, Ref. [76] spread the results of [73] to weakly dependent data and employed their findings to validate the Bayes risk consistency of the relevant discrimination rules. Ref. [77] suggested symmetrized nearest neighbor conditional U -statistics as alternatives to conventional kernel estimators. Ref. [78] took into consideration the functional conditional U -statistic and established the finite-dimensional asymptotic normality. Nevertheless, the non-parametric estimate of the conditional U -statistics in the functional data framework had not received significant attention, despite the subject's relevance. Some recent developments are discussed in references [79,80], in which the authors examine the challenges associated with maintaining a uniform bandwidth in a general framework. The test of independence in the functional framework was based on the Kendall statistics, which may be thought of as examples of the U -statistics; for instance, see [81]. The extension of the investigation described above to conditional empirical U -processes is theoretically attractive, practically helpful, and technically challenging.

The primary objective of this study is to examine a general framework and the weak convergence's characterization of the regular sequence of random spatial functions based on conditional U -processes. This inquiry is simple, as it is difficult to hold the asymptotic equicontinuity under minimal conditions in this general setting, which constitutes a fundamentally unresolved open subject in the literature. We intend to fill this gap in the literature by merging the findings of [37,82,83] with techniques handling the functional data given in [84–87]. However, as demonstrated in the following section, the challenge requires much more than “just” merging concepts from the current outcomes. In fact, complex mathematical derivations are necessary to deal with the typical functional data in our context. This requires the successful application of large-sample theoretical tools, which have been established for empirical processes, where we used the results of the work of [37,82,83].

The structure of the present article is as follows. Section 2 introduces the functional framework and the definitions requested in our work. The assumptions used in our asymptotic analysis go along with a brief discussion. Section 3 gives the uniform rates of the strong convergence. Section 4 includes the paper's main results concerning the uniform weak convergence for the conditional U -processes. In Section 5, we provide some potential applications. In Section 6, we consider the conditional U -statistics in the right censored data framework. In Section 7, we present how to select the bandwidth through the cross-validation procedures. Some concluding remarks and possible future developments are relegated to Section 8. All proofs are gathered in Section 9 to prevent interrupting the presentation flow. Finally, some relevant technical results are given in Appendix A.

2. The Functional Framework

2.1. Notation

For any set $A \subset \mathbb{R}^d$, $|A|$ represents the Lebesgue measure of A and $\llbracket A \rrbracket$ denotes the number of elements in A . For any positive sequence a_n, b_n , we write $a_n \lesssim b_n$ if a constant $C > 0$ independent of n exists such that $a_n \leq Cb_n$ for all n , $a_n \sim b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$, and $a_n \ll b_n$ si $a_n/b_n \rightarrow 0$ as $n \rightarrow \infty$. We use the notation \xrightarrow{d} to indicate convergence in the distribution. We write $X \stackrel{d}{=} Y$ if the random variables X and Y have the same distribution. P_S will denote the joint probability distribution of the sequence of independent and identically distributed (i.i.d.) random vectors $\{S_{0,j}\}_{j \geq 1}$, and $P_{\cdot|S}$ is the conditional probability distribution for $\{S_{0,j}\}_{j \geq 1}$. Let $E_{\cdot|S}$ represent the conditional expectation and $\text{Var}_{\cdot|S}$ represent the variance for $\{S_{0,j}\}_{j \geq 1}$.

2.2. Generality on the Model

In this investigation, we examine the following model:

$$\begin{aligned} &\varphi(\mathbf{Y}_{s_{i_1}, A_n}, \dots, \mathbf{Y}_{s_{i_m}, A_n}) \\ &= r^{(m)} \left(\varphi, \mathbf{X}_{s_{i_1}, A_n}, \dots, \mathbf{X}_{s_{i_m}, A_n}, \frac{s_{i_1}}{A_n}, \dots, \frac{s_{i_m}}{A_n} \right) + \prod_{j=1}^m \sigma \left(\frac{s_{i_j}}{A_n}, \mathbf{x} \right) \epsilon_{i_j} \\ &= r^{(m)} \left(\varphi, \mathbf{X}_{s_{i_1}, A_n}, \dots, \mathbf{X}_{s_{i_m}, A_n}, \frac{s_{i_1}}{A_n}, \dots, \frac{s_{i_m}}{A_n} \right) + \prod_{j=1}^m \epsilon_{s_{i_j}, A_n}, s_{i_j} \in R_n, j = 1, \dots, m, \end{aligned} \tag{4}$$

where $\mathbb{E}[\epsilon_{s, A_n} | \mathbf{X}_{s, A_n}] = 0$ and $R_n = [0, A_n]^d \subset \mathbb{R}^d$ denotes a sampling region with $A_n \rightarrow \infty$ as $n \rightarrow \infty$. Here, \mathbf{Y}_{s, A_n} and \mathbf{X}_{s, A_n} denote random functions in \mathcal{H} and \mathcal{Y} . We consider $\{\mathbf{X}_{s, A_n} : s \in R_n\}$ as a locally stationary random function field on $R_n \subset \mathbb{R}^d$ ($d \geq 2$). As suggested by [88], locally stationary processes are nonstationary time series in which the parameters of the time series can change over time. Locally in time, they can be modeled by a stationary time series, which makes it possible to use asymptotic theories to estimate the parameters of models that depend on time. Time series analyses mostly look at locally stationary models in a parametric framework with coefficients that change over time.

2.3. Local Stationarity

A random function field $\{\mathbf{X}_{s, A_n} : s \in R_n\}$ ($A_n \rightarrow \infty$ as $n \rightarrow \infty$) is considered to be locally stationary if it exhibits behavior that is approximately stationary in the local space. To guaranteed that it is locally stationary around each rescaled space point \mathbf{u} , a process $\{\mathbf{X}_{s, A_n}\}$ can be approximated by a stationary random function field $\{\mathbf{X}_{\mathbf{u}}(\mathbf{s}) : \mathbf{s} \in \mathbb{R}^d\}$ stochastically; for instance, see [89]. The following is one possible way to define this idea.

Definition 1. The \mathcal{H} -valued stochastic process $\{\mathbf{X}_{s, A_n} : s \in R_n\}$ denotes locally stationary if for each rescaled time point $\mathbf{u} \in [0, 1]^d$, there exists an associated \mathcal{H} -valued process $\{\mathbf{X}_{\mathbf{u}}(\mathbf{s}) : \mathbf{s} \in \mathbb{R}^d\}$ with the following properties:

- (i) $\{\mathbf{X}_{\mathbf{u}}(\mathbf{s}) : \mathbf{s} \in \mathbb{R}^d\}$ denotes strictly stationary.
- (ii) It holds that

$$d(\mathbf{X}_{s, A_n}, \mathbf{X}_{\mathbf{u}}(\mathbf{s})) \leq \left(\left\| \frac{\mathbf{s}}{A_n} - \mathbf{u} \right\|_2 + \frac{1}{A_n^d} \right) U_{s, A_n}(\mathbf{u}) \text{ a.s.}, \tag{5}$$

where $\{U_{s, A_n}(\mathbf{u})\}$ denotes a process of positive variables satisfying $\mathbb{E}[(U_{s, A_n}(\mathbf{u}))^\rho] < C$ for some $\rho > 0, C < \infty$; C is independent of \mathbf{u}, \mathbf{s} , and A_n . $\|\cdot\|_2$ is arbitrary norms of \mathbb{R}^d .

The concept of local stationarity for real-valued time series was first presented by [88], and Definition 1 is a natural extension of that idea.

In addition, the definition we offer is the same as that of [90] (Definition 2.1) when \mathcal{H} is the Hilbert space $L^2_{\mathbb{R}}([0, 1])$ of all real-valued functions that are square integrable with respect to the Lebesgue measure on the unit interval $[0, 1]$ with the L_2 -norm given by

$$\|f\|_2 = \sqrt{\langle f, f \rangle}, \langle f, g \rangle = \int_0^1 f(t)g(t)dt,$$

where $f, g \in L^2_{\mathbb{R}}([0, 1])$. In addition to this, the authors provide necessary conditions so that an $L^2_{\mathbb{R}}([0, 1])$ -valued stochastic process $\{X_{t, T}\}$ satisfies (5) with $d(f, g) = \|f - g\|_2$ and $\rho = 2$.

2.4. Sampling Design

We are going to look at the stochastic sampling strategy in order to accommodate the data that are irregularly spaced. First, define R_n as the sampling region. Let $\{A_n\}_{n \geq 1}$ be

a sequence of positive numbers such that $A_n \rightarrow \infty$ as $n \rightarrow \infty$. We consider the sampling region as follows:

$$R_n = [0, A_n]^d. \tag{6}$$

We will discuss the (random) sample designs we will use. Let $f_S(s_0)$ be a continuous, everywhere positive probability density function on $R_0 = [0, 1]^d$, and let $\{S_{0,j}\}_{j \geq 1}$ be a sequence of i.i.d. random vectors with probability density $f_S(s_0)$ such that $\{S_{0,j}\}_{j \geq 1}$ and $\{X_{s,A} : s \in R_n\}$ share a common probability space (Ω, \mathcal{F}, P) and are independent. The realizations $s_{0,1}, \dots, s_{0,n}$ of random vectors $S_{0,1}, \dots, S_{0,n}$ by the following relation:

$$s_j = A_n s_{0,j}, \quad j = 1, \dots, n.$$

gives the sampling sites s_1, \dots, s_n

Herein, we assume that $nA_n^{-d} \rightarrow \infty$ as $n \rightarrow \infty$.

Remark 1. In practice, A_n can be derived by taking the sampling region’s diameter. We can extend the applicability of the assumption (6) to R_n to a broader range of situations, i.e.,

$$R_n = \prod_{j=1}^d [0, A_{j,n}],$$

where $A_{j,n}$ are sequences of positive constants with $A_{j,n} \rightarrow \infty$ as $n \rightarrow \infty$. To avoid more challenging outcomes, we assumed (6). For additional discussion, please refer to [85,87,91–93] and ([94], Chapter 12).

2.5. Mixing Condition

The sequence Z_1, Z_2 , is said to be β -mixing or absolute regular, refer to [95,96], if:

$$\beta(k) := \mathbb{E} \sup_{l \geq 1} \left\{ \left| \mathbb{P}(A | \sigma_l^1) - \mathbb{P}(A) \right| : A \in \sigma_{l+k}^\infty \right\} \rightarrow 0 \text{ as } k \rightarrow \infty.$$

Notably, Ref. [97] produced a comprehensive description of stationary Gaussian processes matching the last condition. Now, we define β -mixing coefficients for a random function field \tilde{X} . Let $\sigma_{\tilde{X}}(T) = \sigma(\{\tilde{X}(s) : s \in T\})$ be the σ -field generated by variables $\{\tilde{X}(s) : s \in T\}$, $T \subset \mathbb{R}^d$. For subsets T_1 and T_2 of \mathbb{R}^d , let

$$\bar{\beta}(T_1, T_2) = \sup \frac{1}{2} \sum_{j=1}^J \sum_{k=1}^K |P(A_j \cap B_k) - P(A_j)P(B_k)|,$$

where the supremum is taken over all pairs of (finite) partitions $\{A_1, \dots, A_J\}$ and $\{B_1, \dots, B_K\}$ of \mathbb{R}^d such that $A_j \in \sigma_{\tilde{X}}(T_1)$ and $B_k \in \sigma_{\tilde{X}}(T_2)$. Furthermore, let

$$d(T_1, T_2) = \inf\{|x - y| : x \in T_1, y \in T_2\},$$

where $|x| = \sum_{j=1}^d |x_j|$ for $x \in \mathbb{R}^d$, and let $\mathcal{R}(b)$ be the collection of all finite disjoint unions of cubes in \mathbb{R}^d with a volume total not exceeding b . Subsequently, the β -mixing coefficients for the random field \tilde{X} can be defined as

$$\beta(a; b) = \sup\{\bar{\beta}(T_1, T_2) : d(T_1, T_2) \geq a, T_1, T_2 \in \mathcal{R}(b)\}. \tag{7}$$

We assume that a non-increasing function β_1 with $\lim_{a \rightarrow \infty} \beta_1(a) = 0$ and a non-decreasing function g_1 exist such that the β -mixing coefficient $\beta(a; b)$ satisfies the following inequality:

$$\beta(a; b) \leq \beta_1(a)g_1(b), \quad a > 0, b > 0, \tag{8}$$

where g_1 may be unbounded for $d \geq 2$.

Remark 2 (Some remarks about mixing conditions). *The size of index sets T_1 and T_2 in the definition of $\beta(a; b)$ must be restricted. Let us explain this point. If the β -mixing coefficients of a random field \mathbf{X} are defined similarly to the β -mixing coefficients for the time series as follows: Let \mathcal{O}_1 and \mathcal{O}_2 be half-planes with boundaries L_1 and L_2 , respectively. For each real number $a > 0$, define*

$$\beta(a) = \sup\{\bar{\beta}(\mathcal{O}_1, \mathcal{O}_2) : d(\mathcal{O}_1, \mathcal{O}_2) \geq a\},$$

where *sup* is taken over all pairs of parallel lines L_1 and L_2 such that $d(L_1, L_2) \geq a$. Subsequently, ([98] Theorem 1) shows that if $\{\mathbf{X}(s) : s \in \mathbb{R}^2\}$ is a strictly stationary mixing random field, and $a > 0$ is a real number. Then, $\beta(a) = 1$ or 0 . This means that if a random field \mathbf{X} is β -mixing ($(\lim_{a \rightarrow \infty} \beta(a) = 0)$), then for η , a positive constant and for some $a > \eta$, the random field \mathbf{X} is “ m -dependent”, i.e., $\beta(a) = 0$. However, this is highly restricted in practice. In order to loosen these results and make them more flexible for practical purposes, it will be necessary to restrict the size of T_1 and T_2 and adopt Definition 7 for the β -mixing. We refer to [85,87,99–101] for additional information on mixing coefficients for random fields.

Ref. [93] writes the form of mixing condition given in Equation (8) for the α -mixing condition and it was considered also in the works of [102,103]. We have considered the β -mixing case, and it is well known that the β -mixing implies the α -mixing. In general, in the expression (8) β_1 is a function defined in a way that it could be dependent on n as the random field \mathbf{X}_{s,A_n} depends on n , yet, g does not, just for the simplicity sake, despite that the general cases where g changes with n are not difficult. We note that the random field \mathbf{Y}_{s,A_n} (or $\varphi(\mathbf{Y}_{s,A_n})$) does not necessarily satisfy the mixing condition (8), since the mixing condition is assumed for \mathbf{X}_{s,A_n} , but with the regression form represented by the model in (4), \mathbf{Y}_{s,A_n} (or $\varphi(\mathbf{Y}_{s,A_n})$) may have a flexible dependence structure.

2.6. Generality on the Model

Let $\{\mathbf{X}_{s,A_n}, \mathbf{Y}_{s,A_n} : s \in R_n\}$ be random variables where \mathbf{Y}_{s,A_n} is in \mathcal{Y} and \mathbf{X}_{s,A_n} takes values in some semi-metric space \mathcal{H} with a semi-metric $d(\cdot, \cdot)$ (a semi-metric (sometimes called pseudo-metric) $d(\cdot, \cdot)$ is a metric which allows $d(x_1, x_2) = 0$ for some $x_1 \neq x_2$) defining a topology to measure the proximity between two elements of \mathcal{H} and which is dissociated from the definition of X in order to prevent concerns with measurability. This study aims to establish the weak convergence of the conditional U -process using the following U -statistic.

$$\begin{aligned} \hat{r}_n^{(m)}(\mathbf{x}, \mathbf{u}; h_n) &:= \hat{r}_n^{(m)}(\varphi, \mathbf{x}, \mathbf{u}; h_n) \\ &= \frac{\sum_{\mathbf{i} \in I_n^m} \varphi(\mathbf{Y}_{s_{i_1}, A_n}, \dots, \mathbf{Y}_{s_{i_m}, A_n}) \prod_{j=1}^m \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}}{\sum_{\mathbf{i} \in I_n^m} \prod_{j=1}^m \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}} \\ &= \frac{\sum_{\mathbf{i} \in I_n^m} \varphi(\mathbf{Y}_{s_{i_1}, A_n}, \dots, \mathbf{Y}_{s_{i_m}, A_n}) \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1 \left(\frac{u_{j,\ell} - \frac{s_{i_j,\ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}}{\sum_{\mathbf{i} \in I_n^m} \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1 \left(\frac{u_{j,\ell} - \frac{s_{i_j,\ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}}, \end{aligned} \tag{9}$$

where

$$I_n^m := \{ \mathbf{i} = (i_1, \dots, i_m) : 1 \leq i_j \leq n \text{ and } i_r \neq i_j \text{ if } r \neq j \},$$

$$\bar{K}(\mathbf{u}) = \prod_{\ell=1}^d K_1(u_\ell), \tag{10}$$

and $\varphi : \mathcal{Y}^m \rightarrow \mathbb{R}$ is a symmetric, measurable function that belongs to some class of functions \mathcal{F}_m , and $\{h_n\}_{n \in \mathbb{N}^*}$ is a sequence of positive real numbers satisfying $h_n \rightarrow 0$ as $n \rightarrow \infty$. In order to examine the weak convergence of the conditional empirical process and the conditional U -process under functional data, we must introduce new notations, let

$$\mathcal{F}_m = \{\varphi : \mathcal{Y}^m \rightarrow \mathbb{R}\},$$

be a point-wise measurable class of real-valued symmetric measurable functions on \mathcal{Y}^m with a measurable envelope function:

$$F(\mathbf{y}) \geq \sup_{\varphi \in \mathcal{F}_m} |\varphi(\mathbf{y})|, \text{ for } \mathbf{y} \in \mathcal{Y}^m. \tag{11}$$

For a kernel function $K(\cdot)$, we define the point-wise measurable class of functions, for $1 \leq m \leq n$

$$\mathcal{K}^m := \left\{ (x_1, \dots, x_m) \mapsto \prod_{i=1}^m K\left(\frac{d(x_i, \cdot)}{h_n}\right), 0 < h_n < 1 \text{ and } (x_1, \dots, x_m) \in \mathcal{K}^m \right\}.$$

We use the notation

$$\psi(\cdot, \cdot) \in \mathcal{F}_m \mathcal{K}^m := \{\varphi_1(\cdot)\varphi_2(\cdot) : \varphi_1 \in \mathcal{F}_m, \varphi_2(\cdot) \in \mathcal{K}^m\},$$

and

$$\psi(\cdot, \cdot) \in \mathcal{F}_1 \mathcal{K}^1 := \mathcal{F} \mathcal{K} = \{\varphi_1(\cdot)\varphi_2(\cdot) : \varphi_1 \in \mathcal{F}_1, \varphi_2(\cdot) \in \mathcal{K}^1\}.$$

2.6.1. Small Ball Probability

In the absence of a universal reference measure, such as the Lebesgue measure, the density function of the functional variable does not exist, which is one of the technical challenges in infinite-dimensional spaces. To circumvent this obstacle, we employ the concept of “small-ball probability”. The function $\phi_x(\cdot)$ precisely controls the concentration of the probability measure of the functional variable on a small ball, which is defined, for a fixed $x \in \mathcal{H}$ for all $r > 0$, by:

$$\mathbb{P}(X \in B(x, r)) =: \phi_x(r) > 0, \tag{12}$$

where the space \mathcal{H} is equipped with the semi-metric $d(\cdot, \cdot)$, and:

$$B(x, r) = \{z \in \mathcal{H} : d(z, x) \leq r\}$$

is a ball in \mathcal{H} with the center x and radius r .

2.6.2. VC-Type Classes of Functions

The asymptotic analysis of functional data is related to concentration properties expressed in terms of the small-ball probability concept. When considering a process indexed by a class of functions, one must also account for other topological concepts, including metric entropy and VC-subgraph classes (“VC” for Vapnik and Červonenkis).

Definition 2. Let $\mathcal{S}_\mathcal{E}$ denote a subset of a semi-metric space \mathcal{E} and N_ϵ a positive integer, a finite set of points $\{e_1, \dots, e_{N_\epsilon}\} \subset \mathcal{E}$ is called, for a given $\epsilon > 0$, a ϵ -net of $\mathcal{S}_\mathcal{E}$ if:

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

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

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

From its name, one can deduce that Kolmogorov invented this idea of metric entropy (cf. Ref. [104]), which was then explored for different metric spaces. This concept was utilized by [105] to provide necessary conditions for the continuity of Gaussian processes. It served as the foundation for remarkable expansions of Donsker's theorem on the weak convergence of empirical processes. $\mathcal{B}_\mathcal{H}$ and $\mathcal{S}_\mathcal{H}$ represent two subsets of the space \mathcal{H} with Kolmogorov's entropy (for the radius ϵ) $\psi_{\mathcal{B}_\mathcal{H}}(\epsilon)$ and $\psi_{\mathcal{S}_\mathcal{H}}(\epsilon)$, respectively, then the Kolmogorov entropy for the subset $\mathcal{B}_\mathcal{H} \times \mathcal{S}_\mathcal{H}$ of the semi-metric space \mathcal{H}^2 by:

$$\psi_{\mathcal{B}_\mathcal{H} \times \mathcal{S}_\mathcal{H}}(\epsilon) = \psi_{\mathcal{B}_\mathcal{H}}(\epsilon) + \psi_{\mathcal{S}_\mathcal{H}}(\epsilon).$$

Hence, $m\psi_{\mathcal{S}_\mathcal{H}}(\epsilon)$ is the Kolmogorov entropy of the subset $\mathcal{S}_\mathcal{H}^m$ of the semi-metric space \mathcal{H}^m . We specify by d the semi-metric on \mathcal{H} ; then, this semi-metric defined on \mathcal{H}^m by:

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

for

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

In this type of study, the semi-metric plays a crucial role. The reader can discover helpful discussions on how to select the semi-metric in [55] (see Chapter 3 and Chapter 13). We must additionally consider another topological term: namely, VC-subgraph classes ("VC" for Vapnik and Červonenkis).

Definition 3. We call a class of subsets \mathcal{C} on a set C a VC-class if there exists a polynomial $\mathfrak{P}(\cdot)$ such that, for every set of N_ϵ points in C , the class \mathcal{C} picks out at most $\mathfrak{P}(N_\epsilon)$ distinct subsets.

Definition 4. A class of functions \mathcal{F} is called a VC-subgraph class if the graphs of the functions in \mathcal{F} form a VC-class of sets, i.e., if we define the subgraph of a real-valued function f on S as the following subset \mathcal{G}_f on $\times \mathbb{R}$:

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

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

A VC-class of functions \mathcal{F} with envelope function F have the following entropy property, for a given $1 \leq q < \infty$, there are constants a and b such as:

$$N(\epsilon, \mathcal{F}, \|\cdot\|_{L_q(Q)}) \leq a \left(\frac{(QF^q)^{1/q}}{\epsilon} \right)^b \tag{13}$$

for any $\epsilon > 0$ and each probability measure such that $QF^q < \infty$. For instance, the following references ([26], Lemma 22), ([106], §4.7), ([107], Theorem 2.6.7), ([108], §9.1) provide a number of sufficient conditions under which (13) holds; refer to ([109], §3.2) for further discussions.

2.7. Conditions and Comments

Assumption 1. (Model and distribution assumptions)

(M1) The \mathcal{H} -valued stochastic process $\{\mathbf{X}_{s,A_n} : s \in \mathbb{R}_n\}$ is locally stationary. Hence, for each time point $\mathbf{u} \in [0, 1]^d$, a strictly stationary process $\{\mathbf{X}_{\mathbf{u}}(s) : s \in \mathbb{R}^d\}$ exists such that for $\|\cdot\|$ an arbitrary norm on \mathbb{R}^d ,

$$d(\mathbf{X}_{s,A_n}, \mathbf{X}_{\mathbf{u}}(s)) \leq \left(\left\| \frac{s}{A_n} - \mathbf{u} \right\|_2 + \frac{1}{A_n^d} \right) U_{s,A_n}(\mathbf{u}) \text{ a.s.}, \tag{14}$$

with $\mathbb{E}[(U_{s,A_n}(\mathbf{u}))^\rho] < C$ for some $\rho \geq 1$ and $C < \infty$ that is independent of \mathbf{u}, s and A_n .

(M2) For $i = 1, \dots, m$, let $B(x_i, h) = \{y \in \mathcal{H} : d(x_i, y) \leq h\}$ be a ball centered at $x_i \in \mathcal{H}$ with radius h , and let $c_d < C_d$ be positive constants. For all $\mathbf{u} \in [0, 1]^d$,

$$\phi_{\mathbf{x}}(h_n) := \mathbb{P}(\mathbf{X}_{\mathbf{u}}(s_1) \in B(x_1, h_n), \dots, \mathbf{X}_{\mathbf{u}}(s_m) \in B(x_m, h_n)) = F_{\mathbf{u}}(h_n, x_1, \dots, x_m)$$

satisfies:

$$0 < c_d \phi(h) f_1(\mathbf{x}) \leq \phi_{\mathbf{x}}(h) \leq C_d \phi(h) f_1(\mathbf{x}), \tag{15}$$

where $\phi(h) \rightarrow 0$ as $h \rightarrow \infty$, and $f_1(\mathbf{x})$ is a non-negative functional in $\mathbf{x} \in \mathcal{H}^m$. Moreover, there exist constants $C_\phi > 0$ and $\varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$,

$$\int_0^\varepsilon \phi(u) du > C_\phi \varepsilon \phi(\varepsilon). \tag{16}$$

(M3) Let $\mathbb{X}_{s,A_n} = (\mathbf{X}_{s_m,A_n}, \dots, \mathbf{X}_{s_1,A_n})$ and $\mathbb{X}_{v,A_n} = (\mathbf{X}_{v_1,A_n}, \dots, \mathbf{X}_{v_m,A_n})$ and $B(\mathbf{x}, h) = \prod_{i=1}^m B(x_i, h)$. Assume

$$\sup_{s,x,A_n} \sup_{s \neq v} \mathbb{P}((\mathbb{X}_{s,A_n}, \mathbb{X}_{v,A_n}) \in B(\mathbf{x}, h) \times B(\mathbf{x}, h)) \leq \psi(h) f_2(\mathbf{x}),$$

where $\psi(h) \rightarrow 0$ as $h \rightarrow 0$, and $f_2(\mathbf{x})$ is a non-negative functional in $\mathbf{x} \in \mathcal{H}^m$. We assume that the ratio $\psi(h) / \phi^2(h)$ is bounded.

(M4) $\sigma : [0, 1] \times \mathcal{H}^m \rightarrow \mathbb{R}$ is bounded by some constant $C_\sigma < \infty$ from above and by some constant $c_\sigma > 0$ from below, that is, $0 < c_\sigma \leq \sigma(\mathbf{u}, \mathbf{x}) \leq C_\sigma < \infty$ for all \mathbf{u} and \mathbf{x} .

(M5) $\sigma(\cdot, \cdot)$ is Lipschitz continuous with respect to \mathbf{u} .

(M6) $\sup_{\mathbf{u} \in [0,1]^m} \sup_{\mathbf{z}: d(\mathbf{x},\mathbf{z}) \leq h} |\sigma(\mathbf{u}, \mathbf{x}) - \sigma(\mathbf{u}, \mathbf{z})| = o(1)$ as $h \rightarrow 0$.

(M7) $r^{(m)}(\mathbf{u}, \mathbf{x})$ is Lipschitz, and it satisfies

$$\sup_{\mathbf{u} \in [0,1]^m} |r^{(m)}(\mathbf{u}_1, \mathbf{x}) - r^{(m)}(\mathbf{u}_2, \mathbf{z})| \leq c_m (d_{\mathcal{H}^m}(\mathbf{x}, \mathbf{z})^\alpha + \|\mathbf{u}_1 - \mathbf{u}_2\|^\alpha) \tag{17}$$

for some $c_m > 0$ and $\alpha > 0$ and the semi-metric $d_{\mathcal{H}^m}(\mathbf{x}, \mathbf{z})$ is defined on \mathcal{H}^m by:

$$d_{\mathcal{H}^m}(\mathbf{x}, \mathbf{z}) := \frac{1}{m} d(x_1, z_1) + \dots + \frac{1}{m} d(x_m, z_m)$$

for $\mathbf{x} = (x_1, \dots, x_m), \mathbf{z} = (z_1, \dots, z_m) \in \mathcal{H}^m$, and it is twice continuously partially differentiable with first derivatives

$$\partial_{u_i} r^{(m)}(\mathbf{u}, \mathbf{x}) = \frac{\partial}{\partial u_i} r^{(m)}(\mathbf{u}, \mathbf{x}),$$

and second derivatives

$$\partial_{u_i u_j}^2 r^{(m)}(\mathbf{u}, \mathbf{x}) = \frac{\partial^2}{\partial u_i \partial u_j} r^{(m)}(\mathbf{u}, \mathbf{x}).$$

Assumption 2. (Kernel assumptions)

- (KB1) The kernel $K_2(\cdot)$ is non-negative, bounded by $\bar{\kappa}$, and has support in $[0, 1]$ such that $0 < K_2(0)$ and $K_2(1) = 0$. Moreover, $K'_2(v) = dK_2(v)/dv$ exists on $[0, 1]$ and satisfies $C'_1 \leq K'_2(v) \leq C'_2$ for two real constants $-\infty < C'_1 < C'_2 < 0$.
- (KB2) The kernel $\bar{K} : \mathbb{R}^d \rightarrow [0, \infty)$ is bounded and has compact support $[-C, C]^d$. Moreover,

$$\int_{[-C, C]^d} \bar{K}(\mathbf{x}) dx = 1, \int_{[-C, C]^d} \mathbf{x}^\alpha \bar{K}(\mathbf{x}) dx = 0, \text{ for any } \alpha \in \mathbb{Z}^d \text{ with } |\alpha| = 1,$$

$$\text{and } |\bar{K}(\mathbf{u}) - \bar{K}(\mathbf{v})| \leq C \|\mathbf{u} - \mathbf{v}\|.$$

- (KB3) The bandwidth h converges to zero at least at a polynomial rate; that is, there exists a small $\xi_1 > 0$ such that $h \leq Cn^{-\xi_1}$ for some constant $0 < C < \infty$.

Assumption 3. (Sampling design assumptions)

- (S1) For any $\alpha \in \mathbb{N}^d$ with $|\alpha| = 1, 2$, $\partial^\alpha f_S(\mathbf{s})$ exists and is continuous on $(0, 1)^d$.
- (S2) $C_0 \leq nA_n^{-d} \leq C_1n^{\eta_1}$ for some $C_0, C_1 > 0$ and small $\eta_1 \in (0, 1)$.

Assumption 4. (Block decomposition assumptions)

- (B1) Let $\{A_{1,n}\}_{n \geq 1}$ and $\{A_{2,n}\}_{n \geq 1}$ be two sequences of positive numbers such that $A_{1,n}, A_{2,n} \rightarrow \infty, A_{2,n} = o(A_{1,n})$, and $A_{1,n} = o(A_n)$, or $\frac{A_{1,n}}{A_n} + \frac{A_{2,n}}{A_{1,n}} \leq C_0^{-1}n^{-\eta} \rightarrow 0$ for some $C_0 > 0$ and $\eta > 0$.
- (B2) We have $\lim_{n \rightarrow \infty} nA_n^{-d} = \kappa \in (0, \infty]$ with $A_n \geq n^{\bar{\kappa}}$ for some $\bar{\kappa} > 0$.
- (B3) We have

$$\left(\frac{1}{nh^{md}\phi(h)}\right)^{1/3} \left(\frac{A_{1,n}}{A_n}\right)^{2d/3} \left(\frac{A_{2,n}}{A_{1,n}}\right)^{2/3} g_1^{1/3}(A_{1,n}^d) \sum_{k=1}^{A_n/A_{1,n}} k^{d-1} \beta_1^{1/3}(kA_{1,n} + A_{2,n}) \rightarrow 0.$$

- (B4) We have $\lim_{n \rightarrow \infty} A_n^d A_{1,n}^{-d} \beta(A_{2,n}; A_n^d) = 0$.

Assumption 5. (Regularity conditions) Let $\alpha_n = \sqrt{\log n / nh^{md}\phi(h)}$. As $n \rightarrow \infty$,

- (R1) $h^{-(md)}\phi(h)^{-1} \alpha_n^{md} A_n^d A_{1,n}^{-d} \beta(A_{2,n}; A_n^d) \rightarrow 0$ and $A_{1,n}^d A_n^{-d} nh^{md}\phi(h)(\log n) \rightarrow 0$,
- (R2) $n^{1/2} h^{(md)/2} \phi(h)^{1/2} / A_{1,n}^d n^{1/\zeta} \geq C_0 n^\eta$ for some $0 < C_0 < \infty$ and $\eta > 0$ and $\zeta > 2$.
- (R3) $A_n^{dp} \phi(h) \rightarrow \infty$, where p is defined in the sequel.

Assumption 6. (E1) For $W_{s_i, A_n} = \prod_{j=1}^m \epsilon_{s_i, A_n}$, it holds that $\sup_{\mathbf{x} \in \mathcal{H}^m} \mathbb{E}|W_{s, A_n}|^\zeta \leq C$ and

$$\sup_{\mathbf{x} \in \mathcal{H}^m} \mathbb{E} \left[|W_{s, A_n}|^\zeta \mid X_{i,n} = \mathbf{x} \right] \leq C$$

for $\zeta > 2$ and $C < \infty$.

- (E2) The β -mixing coefficients of the array $\{\mathbf{X}_{s, A_n}, W_{s, A_n}\}$ satisfy $\beta(a; b) \leq \beta_1(a)g_1(b)$ with $\beta_1(a) \rightarrow 0$ as $a \rightarrow \infty$.

Assumption 7. (Class of functions assumptions)

The classes of functions \mathcal{H}^m and \mathcal{F}_m are such that:

- (C1) The class of functions \mathcal{F}_m is bounded, and its envelope function satisfies for some $0 < M < \infty$:

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

- (C2) The class of functions \mathcal{F}_m is unbounded and its envelope function satisfies for some $p > 2$:

$$\theta_p := \sup_{\mathbf{t} \in \mathcal{S}_{\mathcal{H}}^m} \mathbb{E}(F^p(\mathbf{Y}) \mid \mathbf{x} = \mathbf{x}) < \infty.$$

(C3) The metric entropy of the class $\mathcal{F}_m \mathcal{H}^m$ satisfies, for some $2 < p < \infty$:

$$\begin{aligned} \int_0^\infty (\log N(u, \mathcal{F}_m \mathcal{H}^m, L_1(\mathbb{P}^m)))^{\frac{1}{2}} du &< \infty, \\ \int_0^\infty (\log N(u, \mathcal{F}_m \mathcal{H}^m, L_2(\mathbb{P}^m)))^{\frac{1}{2}} du &< \infty, \\ \int_0^\infty (\log N(u, \mathcal{F}_m \mathcal{H}^m, L_p(\mathbb{P}^m)))^{\frac{1}{2}} du &< \infty. \end{aligned}$$

Comments

When it comes to functional data, traditional statistical methods are entirely ineffective. In our non-parametric functional regression model, we took on the complex theoretical challenge of establishing functional central limit theorems for the conditional U -process under functional absolute regular data and in the context of a two-fold situation. This was accomplished by adopting a two-fold framework. Despite this, the imposed assumptions coincide with some properties of the infinite-dimensional space. These properties include the topological structure on \mathcal{H}^m , the probability distribution of X , and the measurability concept for the classes \mathcal{F}_m and \mathcal{H}^m ; consequently, a discussion regarding the aforementioned assumptions is necessary. The majority of these assumptions were motivated by [37,55,57,84–86,110]. Assumption 1 is the beginning of a formalization of the property of X_i to be locally stationary, and we continue by placing certain restrictions on the distribution behavior of the variables. This allows us to formalize the property in a more precise manner. The condition (M1) refers to the idea of a locally stationary time series, and various random fields can fulfill this requirement. Ref. [85] gave us some examples, and particularly, he proved that this condition is satisfied for locally stationary versions of Lévy-driven moving average random fields. Condition (M2) has been adopted by [84], who in turn was inspired by [57] in his non-parametric density estimation under functional observations. Ref. [84] clarifies that if $\mathcal{H}^m = \mathbb{R}^m$, then the condition overlaps with the fundamental axioms of probability calculus; furthermore, if \mathcal{H}^m is an infinitely dimensional Hilbert space, then $\phi(h_n)$ can drop toward 0 by an exponential speed as $n \rightarrow \infty$. Equation (15) controls the behavior of the small ball probability around zero and is the quite usual condition on the small ball probability. This approximately shows that the small ball probability can be written approximately as the product of two independent functions $\phi^m(\cdot)$ and $f_1(\cdot)$; for instance, for $m = 1$, refer to [111] for the diffusion process, Ref. [112] for a Gaussian measure, Ref. [113] for a general Gaussian process and [84] employed these assumptions for strongly mixing processes. For example, the function $\phi(\cdot)$ can be expressed as $\phi(\epsilon) = \epsilon^\delta \exp(-C/\epsilon^a)$ with $\delta \geq 0$ and $a \geq 0$, and it corresponds to the Ornstein–Uhlenbeck and general diffusion processes (for such processes, $a = 2$ and $\delta = 0$) and the fractal processes (for such processes, $\delta > 0$ and $a = 0$). We refer to the paper of [114] for other examples. Conditions (M4), (M5), (M6) and Assumption 2 represent the regularity conditions, and they are the umbrella that covers the limiting theorems of such a process. Due to the sampling design strategy employed in Section 2.4, a non-uniform density is possible across the sampling region, by which the number of sampling sites is enabled to increase at different rates with respect to the region’s volume $O(A_n^d)$. This sampling design allows the pure increasing domain case ($\lim_{n \rightarrow \infty} nA_n^{-d} = \kappa \in (0, \infty)$) and the mixed increasing domain case ($\lim_{n \rightarrow \infty} nA_n^{-d} = \infty$). Assumption 3 is assumed to address this sampling design and the infill sampling criteria in the stochastic design case, which can also be seen in [94,115]. In addition to the non-uniform possibility of the sampling density, an approach for irregularly spaced sampling sites based on a homogeneous Poisson point process was discussed in ([10], Chapter 8), where the sampling sites must have a uniform distribution over the sampling region. This makes the sampling design used in this work more flexible than the homogeneous Poisson point process and more useful for practical applications. Condition (B1) in Assumption 4 is related to the Blocking technique used to decompose the sampling region R_n into big and small blocks. The sequences $\{A_{1,n}\}$

and $\{A_{2,n}\}$ are related to the large-block–small-block argument, respectively, which is commonly used in proving CLTs for sums of mixing random variables; see [94]. Precisely, $A_{1,n}$ corresponds to the side length of large blocks, while $A_{2,n}$ corresponds to the side length of small blocks. Furthermore, Assumptions 6 help for deriving the weak convergence of the conditional U -statistic $\hat{\psi}$ defined in Section 3. Condition (C1) says that we are dealing with bounded functions, but we are also interested in establishing the functional central limit theorem for conditional U -processes indexed by an unbounded class of functions; therefore, this condition will be replaced by (C2). Each of these generic assumptions is sufficiently weak in connection to the many objects described in our preliminary results. They discuss and utilize the four key axes of this work, which are the topological structure of the functional variables, the probability measure in this functional space, the idea of measurability on the class of functions, and the uniformity governed by the entropy characteristics.

Remark 3. Note that Assumption (C2) in 7 might be substituted by more general hypotheses upon moments of Y as in [109]. That is

(C4)' We denote by $\{\mathcal{M}(x) : x \geq 0\}$ a non-negative continuous function, increasing on $[0, \infty)$, and such that, for some $s > 2$, ultimately as $x \uparrow \infty$,

$$x^{-s} \mathcal{M}(x) \downarrow; \quad x^{-1} \mathcal{M}(x) \uparrow. \tag{18}$$

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

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

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 Convergence Rates for Kernel Estimators

Before expressing the asymptotic behavior of our estimator represented in (9), we will generalize the study to a U -statistic estimator defined by:

$$\hat{\psi}(\mathbf{u}, \mathbf{x}) = \frac{(n-m)!}{n! h^{md} \phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_i, A_n}, \tag{19}$$

where $W_{\mathbf{s}_i, A_n}$ is an array of one-dimensional random variables. In this study, we use the results with $W_{\mathbf{s}_i, A_n} = 1$ and $W_{\mathbf{s}_i, A_n} = \prod_{j=1}^m \epsilon_{\mathbf{s}_{i_j}, A_n}$.

3.1. Hoeffding’s Decomposition

Note $\hat{\psi}(\mathbf{u}, \mathbf{x})$ is a standard U -statistic with a kernel depending on n . We define

$$\begin{aligned} \tilde{\zeta}_j &:= \frac{1}{h^d} \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right), \\ H(Z_1, \dots, Z_m) &:= \prod_{j=1}^m \frac{1}{\phi(h)} K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) W_{\mathbf{s}_i, A_n}, \end{aligned}$$

thus, the U -statistic in (19) can be viewed as a weighted U -statistic of degree m :

$$\hat{\psi}(\mathbf{u}, \mathbf{x}) = \frac{(n-m)!}{n!} \sum_{i \in I_n^m} \tilde{\zeta}_{i_1} \dots \tilde{\zeta}_{i_m} H(Z_{i_1}, \dots, Z_{i_m}). \tag{20}$$

We can write Hoeffding’s decomposition in this case. If we will not assume symmetry for $W_{\mathbf{s}_i, A_n}$ or H , we must define:

- The expectation of $H(Z_{i_1}, \dots, Z_{i_m})$:

$$\theta(i) := \mathbb{E}H(Z_{i_1}, \dots, Z_{i_m}) = \int W_{s_{i,A_n}} \prod_{j=1}^m \frac{1}{\phi(h)} K_2 \left(\frac{d(x_j, v_{s_{i_j,A_n}})}{h_n} \right) d\mathbb{P}_i(z_i). \tag{21}$$

- For all $\ell \in \{1, \dots, m\}$ the position of the argument, construct the function π_ℓ such that:

$$\pi_\ell(z; z_1, \dots, z_{m-1}) := (z_1, \dots, z_{\ell-1}, z, z_\ell, \dots, z_{m-1}).$$

- Define:

$$\begin{aligned} H^{(\ell)}(z; z_1, \dots, z_{m-1}) &:= H\{\pi_\ell(z; z_1, \dots, z_{m-1})\} \\ \theta^{(\ell)}(i; i_1, i_2, \dots, i_{m-1}) &:= \theta\{\pi_\ell(i; i_1, i_2, \dots, i_{m-1})\}. \end{aligned}$$

Hence, the first order expansion of $H(\cdot)$ will be seen as:

$$\begin{aligned} \tilde{H}^{(\ell)}(z) &:= \mathbb{E} \left(H^{(\ell)}(z, Z_1, \dots, Z_{m-1}) \right) \\ &= \int W_{s_{(1, \dots, \ell-1, i, \ell, \dots, m-1), A_n}} \prod_{\substack{j=1 \\ j \neq i}}^{m-1} \frac{1}{\phi(h)} K_2 \left(\frac{d(x_j, v_{s_j, A_n})}{h} \right) \times \frac{1}{\phi(h)} K_2 \left(\frac{d(x_i, v_{s_i, A_n})}{h} \right) \\ &\quad \times \mathbb{P}(dv_1, \dots, dv_{\ell-1}, dv_\ell, \dots, dv_{m-1}) \\ &:= \frac{1}{\phi(h)} K_2 \left(\frac{d(x_i, v_{s_i, A_n})}{h} \right) W_{s_{i, A_n}} \times \int W_{s_{(1, \dots, \ell-1, \ell, \dots, m-1), A_n}} \prod_{\substack{j=1 \\ j \neq i}}^{m-1} \frac{1}{\phi(h)} K_2 \left(\frac{d(x_j, v_{s_j, A_n})}{h} \right) \\ &\quad \times \mathbb{P}(dv_1, \dots, dv_{\ell-1}, dv_\ell, \dots, dv_{m-1}), \end{aligned} \tag{22}$$

with \mathbb{P} as the underlying probability measure, and define

$$f_{i, i_1, \dots, i_{m-1}}^{(\ell)} := \sum_{\ell=1}^m \xi_{i_1} \dots \xi_{i_{\ell-1}} \xi_i \xi_{i_\ell} \dots \xi_{i_{m-1}} \left(\tilde{H}^{(\ell)}(z) - \theta^{(\ell)}(i; i_1, \dots, i_{m-1}) \right). \tag{23}$$

Then, the first-order projection can be defined as:

$$\hat{H}_{1,i}(\mathbf{u}, \mathbf{x}) := \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} f_{i, i_1, \dots, i_{m-1}}^{(\ell)} \tag{24}$$

where

$$I_{n-1}^{m-1}(-i) := \{1 \leq i_1 < \dots < i_{m-1} \leq n \text{ and } i_j \neq i \text{ for all } j \in \{1, \dots, m-1\}\}.$$

For the remainder terms, we denote by $\mathbf{i} \setminus i_\ell := (i_1, \dots, i_{\ell-1}, i_{\ell+1}, \dots, i_m)$ and for $\ell \in \{1, \dots, m\}$, let

$$H_{2,i}(z) := H(z) - \sum_{l=1}^m \tilde{H}_{\mathbf{i} \setminus i_\ell}^{(\ell)}(z_\ell) + (m-1)\theta(\mathbf{i}), \tag{25}$$

where

$$\tilde{H}_{\mathbf{i} \setminus i_\ell}^{(\ell)}(z_\ell) = \mathbb{E}(H(Z_1, \dots, Z_{\ell-1}, z, Z_{\ell+1}, \dots, Z_{m-1})),$$

defined in (22), this projection derives us to the following remainder term:

$$\hat{\psi}_2(\mathbf{u}, \mathbf{x}) := \frac{(n-m)!}{(n)!} \sum_{i \in I_n^m} \xi_{i_1} \dots \xi_{i_m} H_{2,i}(z). \tag{26}$$

Finally, using Equations (24) and (26), and under conditions that:

$$\mathbb{E}(\hat{H}_{1,i}(\mathbf{u}, X)) = 0, \tag{27}$$

$$\mathbb{E}(H_{2,i}(\mathbf{Z} | Z_k)) = 0 \text{ a.s.}, \tag{28}$$

we obtain the [16] decomposition:

$$\begin{aligned} \hat{\psi}(\mathbf{u}, \mathbf{x}) - \mathbb{E}(\hat{\psi}(\mathbf{u}, \mathbf{x})) &= \frac{1}{n} \sum_{i=1}^n \hat{H}_{1,i}(\mathbf{u}, \mathbf{x}) + \hat{\psi}_{2,i}(\mathbf{u}, \mathbf{x}) \\ &:= \hat{\psi}_1(\mathbf{u}, \mathbf{x}) + \hat{\psi}_2(\mathbf{u}, \mathbf{x}). \end{aligned}$$

3.2. Strong Uniform Convergence Rate

We start by giving the following general result concerning the rate of convergence of the U -process presented in (19).

Proposition 1. *Let $\mathcal{F}_m \mathcal{H}^m$ be a measurable VC-subgraph class of functions satisfying Assumption 7 and assume that Assumptions 2, 3, and Condition (B1) in Assumptions 4–6 are also satisfied. Then, the following result holds*

$$\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in [0,1]^m} |\hat{\psi}(\mathbf{u}, \mathbf{x}) - \mathbb{E}[\hat{\psi}(\mathbf{u}, \mathbf{x})]| = O_{P,|S} \left(\sqrt{\frac{\log n}{nh^{md} \phi^m(h)}} \right) P_S - a.s.$$

Next, the uniform rate of convergence of the estimator (9) of the mean function $r^{(m)}$ in the model (4) will be given, using the results of the last proposition.

Theorem 1. *Let $\mathcal{F}_m \mathcal{H}^m$ be a measurable VC-subgraph class of functions satisfying Assumption 7. Let $I_h = [C_1 h, 1 - C_1 h]^{dm}$ and let S_c be a compact subset of \mathcal{H}^m . Suppose that*

$$\inf_{\mathbf{u} \in [0,1]^d} f(\mathbf{u}) > 0.$$

Then, under Assumptions 1–3, Condition (B1) in Assumptions 4–6 (with $W_{s_i, A_n} = 1$ and \mathbf{e}_{s_i, A_n}), the following result holds for P_S almost surely:

$$\begin{aligned} &\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{\mathbf{x} \in S_c} \sup_{\mathbf{u} \in I_h} |\hat{r}_n^{(m)}(\mathbf{x}, \mathbf{u}; h_n) - r^{(m)}(\mathbf{x}, \mathbf{u}; h_n)| \\ &= O_{P,|S} \left(\sqrt{\log n / nh^{md} \phi^m(h)} + h^{2\wedge \alpha} + \frac{1}{A_n^{dp} \phi(h)} \right), \end{aligned}$$

where $p = \min\{1, \rho\}$, and $\rho > 0$ given in Definition 1.

It is worth to note here that the approximation of the functional time series \mathbf{X}_{s, A_n} by a functional stationary random field $\mathbf{X}_u(s)$ provides the error term $A_n^{-dp} \phi^{-1}(h)$.

4. Weak Convergence for Kernel Estimators

In this section, we are interested in studying the weak convergence of the conditional U -process, defined by Equation (9), under absolute regular observations. The following theorem represents the main result in this work concerning the weak convergence of the functional locally stationary random field estimator. Let us define, for $\varphi_1, \varphi_2 \in \mathcal{F}_m$

$$\begin{aligned} \sigma(\varphi_1, \varphi_2) &= E_{,|S} \left(\sqrt{nh^{md} \phi^m(h)} (\hat{r}_n^{(m)}(\varphi_1, \mathbf{x}, \mathbf{u}; h_n) - r^{(m)}(\varphi_1, \mathbf{x}, \mathbf{u})) \right. \\ &\quad \left. \times \sqrt{nh^{md} \phi^m(h)} (\hat{r}_n^{(m)}(\varphi_2, \mathbf{x}, \mathbf{u}; h_n) - r^{(m)}(\varphi_2, \mathbf{x}, \mathbf{u})) \right). \end{aligned} \tag{29}$$

Theorem 2. Let $\mathcal{F}_m \mathcal{H}^m$ be a measurable VC-subgraph class of functions satisfying Assumption 7. Suppose that $f_S(\mathbf{u}) > 0$ and $\epsilon_{s_{i_j}, A_n} = \sigma(s_{i_j} / A_n, \mathbf{x}) \epsilon_{i_j}$, where $\sigma(\cdot, \cdot)$ is continuous and $\{\epsilon_i\}_{i=1}^n$ is a sequence of i.i.d. random variables with mean zero and variance 1. Moreover, suppose $nh^{m(d+1)+4} \rightarrow c_0$ for a constant c_0 . If all assumptions assumed in Theorem 1 hold in addition of Conditions (B2), (B3) and (B4), then the following result holds for P_S almost surely:

$$\sqrt{nh^{md}\phi^m(h)} \left[\hat{r}_n^{(m)}(\varphi, \mathbf{x}, \mathbf{u}; h) - r^{(m)}(\varphi, \mathbf{x}, \mathbf{u}) - B_{\mathbf{u}, \mathbf{x}} \right]$$

converges to a Gaussian process \mathbb{G}_n over $\mathcal{F}_m \mathcal{H}^m$, whose simple paths are bounded and informally continuous with respect to $\|\cdot\|_2$ -norm with co-variance function given by (29) and where the bias term $B_{\mathbf{u}, \mathbf{x}} = O_{P|S}(h^{2\wedge\alpha})$.

Remark 4. Set $A_n^d = O(n^{1-\bar{\eta}_1})$ for some $\bar{\eta}_1 \in [0, 1)$, $A_{1,n} = O(A_n^{\gamma_{A_1}})$, $A_{2,n} = O(A_n^{\gamma_{A_2}})$ with $0 < \gamma_{A_2} < \gamma_{A_1} < 1/3$ and $p = \min\{1, \rho\} = 1$. Assume that we can take a sufficiently large $\zeta > 2$ such that $\frac{2}{\zeta} < (1 - \bar{\eta}_1)(1 - 3\gamma_{A_1})$. Then, Assumption 4 is satisfied for $d \geq 1$.

Remark 5. It is simple to modify the proofs of our results to demonstrate that they still hold when the entropy condition is replaced by the bracketing condition:

$$\int_0^\infty (\log N_{[\cdot]}(u, \mathcal{F}, \mathcal{H}, L_p(\mathbb{P}^m)))^{\frac{1}{2}} du < \infty$$

Refer to p. 270 of [116] for the definition of $N_{[\cdot]}(u, \mathcal{F}, \mathcal{H}, L_p(\mathbb{P}^m))$.

Remark 6. There are basically no restrictions on the choice of the kernel function in our setup apart from satisfying some mild conditions. The selection of the bandwidth, however, is more problematic. It is worth noticing that the choice of the 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 that produces an estimator which has a good balance between the bias and the variance of the considered estimators. It is then more appropriate to consider the bandwidth 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 [117] for more details and discussion on the subject. It would be of interest to establish uniform-in-bandwidth central limit theorems in our setting; i.e., we will let $h > 0$ vary in such a way that $h'_n \leq h \leq h''_n$, where $\{h'_n\}_{n \geq 1}$ and $\{h''_n\}_{n \geq 1}$ are two sequences of positive constants such that $0 < h'_n \leq h''_n < \infty$ and, for either choice of $h_n = h'_n$ or $h_n = h''_n$, fulfills our conditions. It will be of interest to show that

$$\sup_{h'_n \leq h \leq h''_n} \sqrt{nh^{md}\phi^m(h)} \left[\hat{r}_n^{(m)}(\varphi, \mathbf{x}, \mathbf{u}; h) - r^{(m)}(\varphi, \mathbf{x}, \mathbf{u}) - B_{\mathbf{u}, \mathbf{x}} \right]$$

converges to a Gaussian process \mathbb{G}_n over $\mathcal{F}_m \mathcal{H}^m$.

5. Applications

5.1. Metric Learning

Metric learning tries to adapt the metric to the data and has garnered significant interest in recent years; for an overview of metric learning and its applications, see [118,119]. This is prompted by applications ranging from computer vision to bioinformatics-based information retrieval. As an example of the utility of this notion, we give the metric learning issue for supervised classification as described in [119]. Let us consider dependent copies $(\mathbf{X}_{s_1, A_n}, Y_1), \dots, (\mathbf{X}_{s_n, A_n}, Y_n)$ of a $\mathcal{H} \times \mathcal{Y}$ valued random couple (X, Y) , where \mathcal{H} is some feature space and $\mathcal{Y} = \{1, \dots, C\}$, with $C \geq 2$ say, a finite set of labels. Let \mathcal{D} be a set of distance measures $D : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}_+$. The intuitive objective of metric learning in this context is to identify a measure under which points with the same label are close together

and those with different labels are far apart. The standard way to define the risk of a metric D is as follows:

$$R(D) = \mathbb{E}[\phi((1 - D(X, X')) \cdot (2 \mathbf{1}\{Y = Y'\} - 1))], \tag{30}$$

where $\phi(u)$ is a convex loss function upper bounding the indicator function $\mathbf{1}\{u \geq 0\}$: for instance, the hinge loss $\phi(u) = \max(0, 1 - u)$. To estimate $R(D)$, we consider the natural empirical estimator

$$R_n(D) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \bar{K}\left(\frac{\mathbf{u}_i - \mathbf{s}_i / A_n}{h_n}\right) \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_j / A_n}{h_n}\right) \times \phi\left(\left(D(\mathbf{X}_{\mathbf{s}_i, A_n}, \mathbf{X}_{\mathbf{s}_j, A_n}) - 1\right) \cdot (2\pi \{Y_i = Y_j\} - 1)\right), \tag{31}$$

which is one sample U -statistic of degree two with kernel given by:

$$\varphi_D((x, y), (x', y')) = \phi((D(x, x') - 1) \cdot (2 \mathbf{1}\{y = y'\} - 1)).$$

The convergence to (30) of a minimizer of (31), in the non-spatial setting, has been studied in the frameworks of algorithmic stability [120], algorithmic robustness [121] and based on the theory of U -processes under appropriate regularization [122].

5.2. Multipartite Ranking

Let us recall the problem from [119]. Let $X \in \mathcal{H}$ be a random vector of attributes/features and the (temporarily hidden) ordinal labels $Y \in \{1, \dots, K\}$ assigned to it. The goal of multipartite ranking, which uses a training set of labeled examples, is to put the attributes or features in the same order as the labels. Many different fields use this statistical learning problem (e.g., medicine, finance, search engines, e-commerce). Rankings are usually defined by a scoring function, $s : \mathcal{H} \rightarrow \mathbb{R}$, which moves the natural order on the real line to the feature space. The ROC manifold, or its usual summary, the VUS criterion (VUS stands for Volume Under the ROC Surface), is the gold standard for evaluating the ranking performance of $s(x)$; see [123] and the references therein. The best scoring functions, according to [124], are those that are best for all bipartite subproblems. More specifically, they are increasing transformations of the likelihood ratio dF_{k+1}/dF_k , where F_k is the class-conditional distribution for the k th class. When the set of optimal scoring functions is not empty, the authors showed that it is the same as the set of functions that maximizes the amount of space under the ROC surface

$$\text{VUS}(s) = \mathbb{P}\{s(\mathbf{X}_{\mathbf{s}_1}) < \dots < s(\mathbf{X}_{\mathbf{s}_K}) \mid Y_1 = 1, \dots, Y_K = K\}.$$

Given K independent samples $(\mathbf{X}_{\mathbf{s}_1, A_{n_1}}^{(1)}, \dots, \mathbf{X}_{\mathbf{s}_{n_K}, A_{n_K}}^{(K)})$ with distribution $F_k(dx)$ for $k = 1, \dots, K$, the empirical counterpart of the VUS can be written in the following way:

$$\widehat{\text{VUS}}(s) = \frac{1}{\prod_{k=1}^K n_k} \sum_{i_1=1}^{n_1} \dots \sum_{i_k=1}^{n_k} \prod_{j=1}^K \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_{n_j}}{h_n}\right) \mathbf{1}\{s(\mathbf{X}_{\mathbf{s}_{i_1}, A_{n_1}}^{(1)}) < \dots < s(\mathbf{s}_{i_K}, A_{n_K}^{(K)})\}. \tag{32}$$

The empirical VUS (32) is a K -sample U -statistic of degree $(1, \dots, K)$ with kernel given by:

$$\varphi_s(x_1, \dots, x_K) = \mathbf{1}\{s(x_1) < \dots < s(x_K)\}.$$

5.3. Set Indexed Conditional U-Statistics

We aim to study the links between X and Y by estimating functional operators associated with the conditional distribution of Y given X , such as the regression operator, for $C_1 \times \dots \times C_k := \tilde{C}$ in a class of sets \mathcal{C}^m ,

$$\mathbb{G}^{(m)}(C_1 \times \dots \times C_m \mid \mathbf{t}, \mathbf{u}) = \mathbb{E} \left(\prod_{i=1}^m \mathbf{1}\{Y_i \in C_i\} \mid (X_1, \dots, X_k) = (t_1, \dots, t_m) = \mathbf{t} \right) \text{ for } \mathbf{t} \in S_c,$$

where $\mathbf{u} = (u_1, \dots, u_d)$. We define metric entropy with the inclusion of the class of sets \mathcal{C} . For each $\varepsilon > 0$, the covering number is defined as:

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

the quantity $\log(\mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}^{(1)}(\cdot \mid x)))$ is called metric entropy with inclusion of \mathcal{C} with respect to the conditional distribution $\mathbb{G}^{(1)}(\cdot \mid x)$. The quantity $\log \mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}^{(1)}(\cdot \mid x))$ is called metric entropy with inclusion of \mathcal{C} with respect to $\mathbb{G}(\cdot \mid x)$. Estimates for such covering numbers are known for many classes (see, e.g., [125]). We will often assume below that either $\log \mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}^{(1)}(\cdot \mid x))$ or $\mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}^{(1)}(\cdot \mid x))$ behave like powers of ε^{-1} : we say that the condition (R_γ) holds if

$$\log \mathcal{N}(\varepsilon, \mathcal{C}, \mathbb{G}^{(1)}(\cdot \mid x)) \leq H_\gamma(\varepsilon), \text{ for all } \varepsilon > 0, \tag{33}$$

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 [126], it is worth noticing that the condition (33), $\gamma = 0$, holds for intervals, rectangles, balls, ellipsoids, and for classes which are constructed from the above by performing set operations union, intersection and complement finitely many times. The classes of convex sets in \mathbb{R}^d ($d \geq 2$) fulfill the condition (33), $\gamma = (d - 1)/2$. This and other classes of sets satisfying (33) with $\gamma > 0$ can be found in [125]. As a particular case of (9), we estimate $\mathbb{G}^{(m)}(C_1 \times \dots \times C_m \mid \mathbf{t})$

$$\widehat{\mathbb{G}}_n^{(m)}(\tilde{C}, \mathbf{t}, \mathbf{u}) = \frac{\sum_{i \in I_n^m} \prod_{j=1}^m \mathbf{1}\{Y_{s_{ij}, A_n} \in C_j\} \left\{ \prod_{\ell=1}^d K_1 \left(\frac{u_{j,\ell} - \frac{s_{ij,\ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x_j, X_{s_{ij}, A_n})}{h_n} \right) \right\}}{\sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1 \left(\frac{u_{j,\ell} - \frac{s_{ij,\ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x_j, X_{s_{ij}, A_n})}{h_n} \right) \right\}}. \tag{34}$$

One can apply Theorem 1 to infer that, in probability,

$$\sup_{\tilde{C} \in \mathcal{C}^m} \sup_{\mathbf{t} \in S_c, \mathbf{u} \in I_h} \left| \widehat{\mathbb{G}}_n^{(m)}(\tilde{C}, \mathbf{t}) - \mathbb{G}^{(m)}(\tilde{C} \mid \mathbf{t}) \right| \rightarrow 0. \tag{35}$$

Remark 7. Another point of view is to consider the following situation, for a compact $J \subset \mathbb{R}^{dm}$,

$$\mathbf{G}^{(m)}(\mathbf{y}_1, \dots, \mathbf{y}_k \mid \mathbf{t}, \mathbf{u}) = \mathbb{E} \left(\prod_{i=1}^m \mathbf{1}\{Y_i \leq \mathbf{y}_i\} \mid (X_1, \dots, X_m) = \mathbf{t} \right) \text{ for } \mathbf{t} \in S_c, (\mathbf{y}_1, \dots, \mathbf{y}_m) \in J.$$

Let $L(\cdot)$ be a distribution in \mathbb{R}^d and h_n is a sequence of positive real numbers. One can estimate $\mathbf{G}^{(m)}(\mathbf{y}_1, \dots, \mathbf{y}_m \mid \mathbf{t}, \mathbf{u}) = \mathbf{G}^{(m)}(\mathbf{y} \mid \mathbf{t}, \mathbf{u})$ by

$$\widehat{\mathbf{G}}_n^{(m)}(\mathbf{y}, \mathbf{t}, \mathbf{u}) := \frac{\sum_{i \in I_n^m} \prod_{j=1}^m L\left(\frac{\mathbf{Y}_{s_{ij}, A_n} - \mathbf{t}_j}{h_n}\right) \left\{ \prod_{\ell=1}^d K_1\left(\frac{u_{j,\ell} - \frac{s_{ij,\ell}}{A_n}}{h_n}\right) K_2\left(\frac{d(x_j, \mathbf{X}_{s_{ij}, A_n})}{h_n}\right) \right\}}{\sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1\left(\frac{u_{j,\ell} - \frac{s_{ij,\ell}}{A_n}}{h_n}\right) K_2\left(\frac{d(x_j, \mathbf{X}_{s_{ij}, A_n})}{h_n}\right) \right\}}$$

One can use Theorem 1 to infer that, in probability,

$$\sup_{\mathbf{t} \in S_c} \sup_{\mathbf{u} \in I_h} \sup_{\mathbf{y} \in J} \left| \widehat{\mathbf{G}}_n^{(m)}(\mathbf{y}, \mathbf{t}, \mathbf{u}) - \mathbf{G}^{(m)}(\mathbf{y} \mid \mathbf{t}, \mathbf{u}) \right| \rightarrow 0. \tag{36}$$

5.4. Discrimination

Now, we apply the results on the problem of discrimination described in Section 3 of [127], refer to also [128]. We will use similar notations and settings. Let $\varphi(\cdot)$ be any function taking at most finitely many values, say $1, \dots, M$. The sets

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

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

$$\mathbb{P}(g(\mathbf{X}_1, \dots, \mathbf{X}_m) = \varphi(\mathbf{Y}_1, \dots, \mathbf{Y}_m)) \leq \sum_{j=1}^M \int_{\{x: g(\mathbf{t})=j\}} \max m^j(\mathbf{t}) d\mathbb{P}(\mathbf{t}),$$

where

$$m^j(\mathbf{t}) = \mathbb{P}(\varphi(\mathbf{Y}_1, \dots, \mathbf{Y}_m) = j \mid \mathbf{X}_1, \dots, \mathbf{X}_m = \mathbf{t}), \quad \mathbf{t} \in S_c.$$

The above inequality becomes equality if

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

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

$$L^* = 1 - \mathbb{P}(g_0(\mathbf{X}_1, \dots, \mathbf{X}_m) = \varphi(\mathbf{Y}_1, \dots, \mathbf{Y}_m)) = 1 - \mathbb{E} \left\{ \max_{1 \leq j \leq M} m^j(\mathbf{t}) \right\}$$

is called the Bayes risk. Each of the above unknown functions $m^j(\cdot)$ values can be consistently estimated by one of the methods discussed in the preceding sections. Let, for $1 \leq j \leq M$,

$$m_n^j(x, \mathbf{u}) = \frac{\sum_{i \in I_n^m} \mathbf{1}\{\varphi(\mathbf{Y}_{s_{i1}, A_n}, \dots, \mathbf{Y}_{s_{im}, A_n}) = j\} \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1\left(\frac{u_{j,\ell} - \frac{s_{ij,\ell}}{A_n}}{h_n}\right) K_2\left(\frac{d(x_j, \mathbf{X}_{s_{ij}, A_n})}{h_n}\right) \right\}}{\sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1\left(\frac{u_{j,\ell} - \frac{s_{ij,\ell}}{A_n}}{h_n}\right) K_2\left(\frac{d(x_j, \mathbf{X}_{s_{ij}, A_n})}{h_n}\right) \right\}}$$

Set

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

Let us introduce

$$L_n^* = \mathbb{P}(g_{0,n}(X_1, \dots, X_m) \neq \varphi(Y_1, \dots, Y_m)).$$

Then, one can show that the discrimination rule $g_{0,n}(\cdot)$ is asymptotically Bayes' risk consistent

$$L_n^* \rightarrow L^*.$$

This follows from the obvious relation

$$|L^* - L_n^*| \leq 2\mathbb{E} \left[\max_{1 \leq j \leq M} |m_n^j(\mathbf{X}, \mathbf{u}) - m^j(\mathbf{X})| \right].$$

6. Extension to the Censored Case

Consider a triple (Y, C, X) of random variables defined in $\mathbb{R} \times \mathbb{R} \times \mathcal{H}$. Here, Y is the variable of interest, C is a censoring variable and X is a concomitant variable. Throughout, we will use [129] notation and we work with a sample $\{(Y_i, C_i, X_{s_i, A_n})\}$ of identically distributed replication of (Y, C, X) , $n \geq 1$. Actually, in the right censorship model, the pairs (Y_i, C_i) , $1 \leq i \leq n$, are not directly observed, and the corresponding information is given by $Z_i := \min\{Y_i, C_i\}$ and $\delta_i := \mathbf{1}\{Y_i \leq C_i\}$, $1 \leq i \leq n$. Accordingly, the observed sample is

$$\mathcal{D}_n = \{(Z_i, \delta_i, X_{s_i, A_n}), i = 1, \dots, n\}.$$

Survival data in clinical trials or failure time data in reliability studies, for example, are often subject to such censoring. More specifically, many statistical experiments result in incomplete samples, even under well-controlled conditions. For example, clinical data for surviving most types of disease are usually censored by other competing risks to life which result in death. In the sequel, we impose the following assumptions upon the distribution of (X, Y) . For $-\infty < t < \infty$, set

$$F_Y(t) = \mathbb{P}(Y \leq t), \quad G(t) = \mathbb{P}(C \leq t), \quad \text{and} \quad H(t) = \mathbb{P}(Z \leq t),$$

the right-continuous distribution functions of Y , C and Z respectively. For any right-continuous distribution function L defined on \mathbb{R} , denote by

$$T_L = \sup\{t \in \mathbb{R} : L(t) < 1\}$$

the upper point of the corresponding distribution. Now, consider a point-wise measurable class \mathcal{F} of real measurable functions defined on \mathbb{R} , and assume that \mathcal{F} is of VC-type. We recall the regression function of $\psi(Y)$ evaluated at $X = x$, for $\psi \in \mathcal{F}$ and $x \in \mathcal{H}$, given by

$$r^{(1)}(\psi, x) = \mathbb{E}(\psi(Y) | X = x),$$

when Y is right-censored. To estimate $r^{(1)}(\psi, \cdot)$, we make use of the Inverse Probability of Censoring Weighted (I.P.C.W.) estimators have recently gained popularity in the censored data literature (see [130–132]). The key idea of I.P.C.W. estimators is as follows. Introduce the real-valued function $\Phi_\psi(\cdot, \cdot)$ defined on \mathbb{R}^2 by

$$\Phi_\psi(y, c) = \frac{\mathbf{1}\{y \leq c\} \psi(y \wedge c)}{1 - G(y \wedge c)}. \tag{37}$$

Assuming the function $G(\cdot)$ to be known, first note that $\Phi_\psi(Y_i, C_i) = \delta_i \psi(Z_i) / (1 - G(Z_i))$ is observed for every $1 \leq i \leq n$. Moreover, under the Assumption (\mathcal{S}) below,

(\mathcal{S}) C and (Y, X) are independent.

We have

$$r^{(1)}(\Phi_\psi, x) := \mathbb{E}(\Phi_\psi(Y, C) | X = x)$$

$$\begin{aligned}
 &= \mathbb{E} \left\{ \frac{\mathbf{1}\{Y \leq C\} \psi(Z)}{1 - G(Z)} \mid \mathbb{X} = \mathbf{x} \right\} \\
 &= \mathbb{E} \left\{ \frac{\psi(Y)}{1 - G(Y)} \mathbb{E}(\mathbf{1}\{Y \leq C\} \mid \mathbf{X}, Y) \mid \mathbf{X} = \mathbf{x} \right\} \\
 &= r^{(1)}(\psi, \mathbf{x}).
 \end{aligned} \tag{38}$$

Therefore, any estimate of $r^{(1)}(\Phi_\psi, \cdot)$, which can be built on fully observed data, turns out to be an estimate for $r^{(1)}(\psi, \cdot)$ too. Thanks to this property, most statistical procedures known to provide estimates of the regression function in the uncensored case can be naturally extended to the censored case. For instance, kernel-type estimates are particularly easy to construct. Set, for $x \in \mathcal{H}$, $h \geq l_n$, $1 \leq i \leq n$,

$$\bar{\omega}_{n, K_{1,2}, h_{n,i}}^{(1)}(x, u) := \frac{\prod_{\ell=1}^d K_1 \left(\frac{u_\ell - \frac{s_{j,\ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x, \mathbf{X}_{s_j, A_n})}{h_n} \right)}{\sum_{j=1}^n \prod_{\ell=1}^d K_1 \left(\frac{u_\ell - \frac{s_{j,\ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x, \mathbf{X}_{s_j, A_n})}{h_n} \right)}. \tag{39}$$

In view of (37)–(39), whenever $G(\cdot)$ is known, a kernel estimator of $r^{(1)}(\psi, \cdot)$ is given by

$$\check{r}_n^{(1)}(\psi, x, u; h_n) = \sum_{i=1}^n \bar{\omega}_{n, K_{1,2}, h_{n,i}}^{(1)}(x, u) \frac{\delta_i \psi(Z_i)}{1 - G(Z_i)}. \tag{40}$$

The distribution function $G(\cdot)$ is generally unknown and has to be estimated. We will denote by $G_n^*(\cdot)$ the Kaplan–Meier estimator of the function $G(\cdot)$ [133]. Namely, adopting the conventions

$$\prod_{\emptyset} = 1$$

and $0^0 = 1$ and setting

$$N_n(u) = \sum_{i=1}^n \mathbf{1}\{Z_i \geq u\},$$

we have

$$G_n^*(u) = 1 - \prod_{i: Z_i \leq u} \left\{ \frac{N_n(Z_i) - 1}{N_n(Z_i)} \right\}^{(1-\delta_i)}, \text{ for } u \in \mathbb{R}.$$

Given this notation, we will investigate the following estimator of $r^{(1)}(\psi, \cdot)$

$$\check{r}_n^{(1)*}(\psi, x, u; h_n) = \sum_{i=1}^n \bar{\omega}_{n, K_{1,2}, h_{n,i}}^{(1)}(x, u) \frac{\delta_i \psi(Z_i)}{1 - G_n^*(Z_i)}, \tag{41}$$

refer to [129,130]. Adopting the convention $0/0 = 0$, this quantity is well defined, since $G_n^*(Z_i) = 1$ if and only if $Z_i = Z_{(n)}$ and $\delta_{(n)} = 0$, where $Z_{(k)}$ is the k th-ordered statistic associated with the sample (Z_1, \dots, Z_n) for $k = 1, \dots, n$ and $\delta_{(k)}$ is the δ_j corresponding to $Z_k = Z_j$. When the variable of interest is right-censored, the functional of the (conditional) law can generally not be estimated on the complete support (see [132]). To obtain our results, we will work under the following assumptions.

(A.1) $\mathcal{F} = \{\psi := \psi_1 \mathbf{1}\{(-\infty, \tau)^m\}, \psi_1 \in \mathcal{F}_1\}$, where $\tau < T_H$ and \mathcal{F}_1 is a point-wise measurable class of real measurable functions defined on \mathbb{R} and of type VC.

(A.2) The class of functions \mathcal{F} has a measurable and uniformly bounded envelope function Y with,

$$Y(y_1, \dots, y_k) \geq \sup_{\psi \in \mathcal{F}} |\psi(y_1, \dots, y_k)|, \quad y_i \leq T_H.$$

We now have all the ingredients to state the result corresponding to the censored case. By combining the results of Proposition 9.6 and Lemma 9.7 of [134], Theorem 1, we have, in probability,

$$\sup_{x,u} \left| \check{r}_n^{(1)*}(\psi, x, u; h_n) - \widehat{\mathbb{E}}(\check{r}_n^{(1)*}(\psi, x, u; h_n)) \right| \rightarrow 0. \tag{42}$$

A right-censored version of an unconditional U -statistic with a kernel of degree $m \geq 1$ is introduced by the principle of a mean preserving reweighting scheme in [135]. Ref. [136] has proved almost sure convergence of multi-sample U -statistics under random censorship and provided application by considering the consistency of a new class of tests designed for testing equality in distribution. To overcome potential biases arising from right-censoring of the outcomes and the presence of confounding covariates, Ref. [137] proposed adjustments to the classical U -statistics. Ref. [138] proposed a different way in the estimation procedure of the U -statistic by using a substitution estimator of the conditional kernel given the observed data. To our best knowledge, the problem of the estimation of the conditional U -statistics was opened up to the present, and it gives our main motivation to the study of this section. A natural extension of the function defined in (37) is given by

$$\Phi_\psi(y_1, \dots, y_m, c_1, \dots, c_m) = \frac{\prod_{i=1}^m \{\mathbf{1}\{y_i \leq c_i\} \psi(y_1 \wedge c_1, \dots, y_m \wedge c_m)\}}{\prod_{i=1}^m \{1 - G(y_i \wedge c_i)\}}.$$

From this, we have an analogous relation to (38) given by

$$\begin{aligned} & \mathbb{E}(\Phi_\psi(Y_1, \dots, Y_m, C_1, \dots, C_m) \mid (\mathbf{X}_1, \dots, \mathbf{X}_m) = \mathbf{t}) \\ &= \mathbb{E}\left(\frac{\prod_{i=1}^m \{\mathbf{1}\{Y_i \leq C_i\} \psi(Y_1 \wedge C_1, \dots, Y_m \wedge C_m)\}}{\prod_{i=1}^m \{1 - G(Y_i \wedge C_i)\}} \mid (\mathbf{X}_1, \dots, \mathbf{X}_m) = \mathbf{t}\right) \\ &= \mathbb{E}\left(\frac{\psi(Y_1, \dots, Y_m)}{\prod_{i=1}^m \{1 - G(Y_i)\}} \mathbb{E}\left(\prod_{i=1}^m \{\mathbf{1}\{Y_i \leq C_i\} \mid (Y_1, \mathbf{X}_1), \dots, (Y_m, \mathbf{X}_m)\right) \mid (\mathbf{X}_1, \dots, \mathbf{X}_m) = \mathbf{t}\right) \\ &= \mathbb{E}(\psi(Y_1, \dots, Y_m) \mid (\mathbf{X}_1, \dots, \mathbf{X}_m) = \mathbf{t}) = m_\psi(\mathbf{t}). \end{aligned}$$

An analogue estimator to (9) in the censored case is given by

$$\check{r}_n^{(m)}(\psi, \mathbf{t}, \mathbf{u}; h_n) = \sum_{(i_1, \dots, i_m) \in I(m, n)} \frac{\delta_{i_1} \dots \delta_{i_m} \psi(Z_{i_1}, \dots, Z_{i_m})}{(1 - G(Z_{i_1})) \dots (1 - G(Z_{i_m}))} \bar{\omega}_{n, K_{1,2}, h_n, i}^{(m)}(\mathbf{t}, \mathbf{u}), \tag{43}$$

where, for $\mathbf{i} = (i_1, \dots, i_k) \in I(k, n)$,

$$\bar{\omega}_{n, K_{1,2}, h_n, i}^{(k)}(\mathbf{x}, \mathbf{u}) = \frac{\prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1\left(\frac{u_{j,\ell} - \frac{s_{i_j, \ell}}{A_n}}{h_n}\right) K_2\left(\frac{d(x_j, \mathbf{X}_{s_{i_j, A_n}})}{h_n}\right) \right\}}{\sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1\left(\frac{u_{j,\ell} - \frac{s_{i_j, \ell}}{A_n}}{h_n}\right) K_2\left(\frac{d(x_j, \mathbf{X}_{s_{i_j, A_n}})}{h_n}\right) \right\}}. \tag{44}$$

The estimator that we will investigate is given by

$$\check{r}_n^{(m)*}(\psi, \mathbf{t}, \mathbf{u}; h_n) = \sum_{(i_1, \dots, i_k) \in I(m, n)} \frac{\delta_{i_1} \dots \delta_{i_m} \psi(Z_{i_1}, \dots, Z_{i_m})}{(1 - G_n^*(Z_{i_1})) \dots (1 - G_n^*(Z_{i_m}))} \bar{\omega}_{n, K_{1,2}, h_n, i}^{(k)}(\mathbf{t}, \mathbf{u}). \tag{45}$$

Corollary 1. Under the assumptions (A.1)–(A.2) and the conditions of Theorem 1, we have

$$\sup_{x \in \mathcal{H}^m} \sup_{\mathbf{u} \in I_n, x \in S_c} \left| \check{r}_n^{(m)*}(\psi, \mathbf{t}, \mathbf{u}; h_n) - \mathbb{E}\check{r}_n^{(m)*}(\psi, \mathbf{t}, \mathbf{u}; h_n) \right|$$

$$= O_{P,|s} \left(\sqrt{\log n / nh^{md} \phi^m(h)} + \frac{1}{A_n^{dp} \phi(h)} \right),$$

In the last corollary, we use the law of iterated logarithm for $G_n^*(\cdot)$ established in [139] ensuring that

$$\sup_{t \leq \tau} |G_n^* - G(t)| = O \left(\sqrt{\frac{\log \log n}{n}} \right) \text{ almost surely as } n \rightarrow \infty. \tag{46}$$

At this point, we may refer to [69,134,140].

7. The Bandwidth Selection Criterion

Many methods have been established and developed to construct, in asymptotically optimal ways, bandwidth selection rules for non-parametric kernel estimators especially for the Nadaraya–Watson regression estimator we quote among them [141–143]. This parameter has to be selected suitably, either in the standard finite dimensional case, or in the infinite dimensional framework for ensuring good practical performances. However, according to our knowledge, such studies do not presently exist for treating a such general functional conditional U -statistic. Nevertheless, an extension of the leave-one-out cross-validation procedure allows us to define, for any fixed $\mathbf{j} = (j_1, \dots, j_m) \in I_n^m$:

$$\begin{aligned} & \widehat{r}_{n,j}^{(m)}(\mathbf{x}, \mathbf{u}; h_n) \\ &= \frac{\sum_{i \in I_n^m(j)} \varphi(\mathbf{Y}_{s_{i_1, A_n}}, \dots, \mathbf{Y}_{s_{i_m, A_n}}) \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1 \left(\frac{u_{j,\ell} - \frac{s_{i_j, \ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j, A_n}})}{h_n} \right) \right\}}{\sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \prod_{\ell=1}^d K_1 \left(\frac{u_{j,\ell} - \frac{s_{i_j, \ell}}{A_n}}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j, A_n}})}{h_n} \right) \right\}}, \tag{47} \end{aligned}$$

where

$$I_n^m(j) := \{i \in I_n^m \text{ and } i \neq j\} = I_n^m \setminus \{j\}.$$

The Equation (47) represents the leave-one-out- $(\mathbf{X}_j, \mathbf{Y}_j)$ estimator of the functional regression and also could be considered as a predictor of $\varphi(\mathbf{Y}_{s_{j_1, A_n}}, \dots, \mathbf{Y}_{s_{j_m, A_n}}) := \varphi(\mathbf{Y}_j)$. In order to minimize the quadratic loss function, we introduce the following criterion: we have for some (known) non-negative weight function $\mathscr{W}(\cdot)$:

$$CV(\varphi, h_n) := \frac{(n-m)!}{n!} \sum_{j \in I_n^m} \left(\varphi(\mathbf{Y}_j) - \widehat{r}_{n,j}^{(m)}(\mathbf{X}_j, \mathbf{u}; h_n) \right)^2 \widetilde{\mathscr{W}}(\mathbf{X}_j), \tag{48}$$

where $\mathbf{X}_j = (\mathbf{X}_{s_{j_1, A_n}}, \dots, \mathbf{X}_{s_{j_m, A_n}})$. Following the ideas developed by [143], a natural way for choosing the bandwidth is to minimize the precedent criterion, so let us choose $\widehat{h}_n \in [a_n, b_n]$ minimizing among $h \in [a_n, b_n]$:

$$CV(\varphi, h_n).$$

The main interest of our results is the possibility to derive the asymptotic properties of our estimate even if the bandwidth parameter is a random variable, as in the last equation. Following [144] where the bandwidths are locally chosen by a data-driven method based on the minimization of a functional version of a cross-validated criterion, one can replace (48) by

$$CV(\varphi, h_n) := \frac{(n-m)!}{n!} \sum_{j \in I_n^m} \left(\varphi(\mathbf{Y}_j) - \widehat{r}_{n,j}^{(m)}(\mathbf{X}_j, \mathbf{u}; h_n) \right)^2 \widehat{\mathscr{W}}(\mathbf{X}_j, \mathbf{x}), \tag{49}$$

where

$$\widehat{\mathcal{W}}(\mathbf{s}, \mathbf{t}) := \prod_{i=1}^m \widehat{W}(\mathbf{s}_i, \mathbf{t}_i).$$

In practice, one takes for $\mathbf{i} \in I_n^m$ the uniform global weights $\widehat{\mathcal{W}}(\mathbf{X}_i) = 1$, and the local weights

$$\widehat{W}(\mathbf{X}_i, \mathbf{t}) = \begin{cases} 1 & \text{if } d(\mathbf{X}_i, \mathbf{t}) \leq h, \\ 0 & \text{otherwise.} \end{cases}$$

For the sake of brevity, we have just considered the most popular method: that is, the cross-validated selected bandwidth. This may be extended to any other bandwidth selector such as the bandwidth based on Bayesian ideas [145].

Remark 8. For notational convenience, we have chosen the same bandwidth sequence for each margin. This assumption can be dropped easily if one wants to use the vector bandwidths (see, in particular, Chapter 12 of [6]). With obvious changes of notation, our results and their proofs remain true when h_n is replaced by a vector bandwidth $\mathbf{h}_n = (h_n^{(1)}, \dots, h_n^{(d)})$, where $\min h_n^{(i)} > 0$. In this situation, we set $h_n = \prod_{i=1}^d h_n^{(i)}$, and for any vector $\mathbf{v} = (v_1, \dots, v_d)$, we replace \mathbf{v}/h_n by $(v_1/h_n^{(1)}, \dots, v_d/h_n^{(d)})$. For ease of presentation, we chose to use real-valued bandwidths throughout.

Remark 9. We mention that a different bandwidth criterion suggested by [1] is the rule of thumb. Strictly speaking, since the cross-validated bandwidth is random, the asymptotic theory can only be justified with this random bandwidth via a specific stochastic equicontinuity argument. Cross-validation is employed by [146] to examine the equality of two unconditional and conditional functions in the context of mixed categorical and continuous data. However, this approach, which is optimal for estimation, loses its optimality when applied to non-parametric kernel testing. For testing a parametric model for conditional mean function against a non-parametric alternative, Ref. [147] proposed an adaptive-rate-optimal rule. Ref. [148] present the other method for selecting a proper bandwidth. Ref. [148] propose, utilizing the Edgeworth expansion of the asymptotic distribution of the test, to select the bandwidth such that the power function of the test is maximized while the size function is controlled. Future investigation will focus on the aforementioned three approaches.

8. Concluding Remarks

In this paper, we considered the kernel type estimator for conditional U -statistics, including a particular case, the Nadaraya–Watson estimator, in a functional setting with random fields. To obtain our results, we ought to make assumptions requiring some regularity on the conditional U -statistics and conditional moments, some decay rates on the probability of the variables belonging to shrinking open balls, and suitable decreasing rates on the mixing coefficients. Mainly, the conditional moment assumption enables the consideration of unbounded classes of functions. The proof of the weak convergence respects a typical technique: finite dimensional convergence and equicontinuity of the conditional U -processes.

Both results, the uniform rate of convergence and the weak convergence, are grounded on a general blocking technique adjusted for irregularly spaced sampling sites, where we need to pay attention to the effect of the non-equidistant sampling sites. We intricately reduce the work to the independent setting to address this issue. Indeed, as there is no practical guidance for introducing order to spatial points as opposed to time series, not asymptotically but exactly independent blocks of observations have been constructed by ([149], Corollary 2.7) (Lemma A4) and then results of independent data could be applied directly to the independent blocks. Here, Ref. [149] declares that the uniform convergence result requires the β -mixing condition to connect the original sequence with the sequence of the independent blocks, and this connection still holds under the ϕ -mixing condition but

is not necessary under the α -mixing conditions. Therefore, we use the β -mixing sequence as we aim to derive the weak convergence for processes indexed by classes of functions.

Ref. [85] in his work gives us a possible extension of the sampling region inspired by [93]. This extension can be explained as follows. It is feasible to generalize the definition of the sample region R_n to include non-standard forms. For instance, we may use the sample region concept [93] as follows: First, let R_n be the sampling region. Define R_0^* as an open connected subset of $(-2, 2)^d$ containing $[-1, 1]^d$ and R_0 as a Borel set such that $R_0^* \subset R_0 \subset \bar{R}_0^*$, and where for any set $S \subset \mathbb{R}^d$, \bar{S} signifies its closure. Let $\{A_n\}_{n \geq 1}$ be a sequence of positive numbers such that $A_n \rightarrow \infty$ as $n \rightarrow \infty$ and define $R_n = A_n R_0$ as a sampling region. In addition, for any sequence of positive numbers $\{a_n\}_{n \geq 1}$ with $a_n \rightarrow 0$ as $n \rightarrow \infty$, let $O(a_n^{-d+1})$, as $n \rightarrow \infty$, be the number of cubes of the form $a_n(i + [0, 1]^d)$, $i \in \mathbb{Z}^d$ with their lower left corner $a_n i$ on the lattice $a_n \mathbb{Z}^d$ that intersects both R_0 and R_0^c (see Condition B in [93], Chapter 12, Section 12.2) (This condition is the prototype R_0 boundary's condition; it must always be assumed on the region R_n to prevent pathological situations, and it is satisfied by the majority of areas of practical significance. This condition is satisfied in the plane ($d = 2$), for instance, if the boundary ∂R_0 of R_0 is defined by a simple rectifiable curve of limited length. When sample sites are defined on the integer grid \mathbb{Z}^d , this condition means that the effect of data points toward the boundary of R_n is small compared to the overall number of data points). In addition, define f as a continuous, everywhere positive probability density function on R_0 , and let $\{S_{0,i}\}_{i \geq 1}$ be a sequence of i.i.d. random vectors with density f . Assume that $\{S_{0,i}\}_{i \geq 1}$ and X_{s,A_n} are independent. Replacing our setting in Section 2.4 with this new one, our results still hold, and it will be possible to show uniform convergence and weak convergence under the same assumptions and identical proofs. For future investigation, it will be interesting to relax the mixing conditions to the weak dependence (or the ergodicity framework). This generalization is nontrivial, since we need some maximal moment inequalities in our asymptotic results that are not available in this setting. Another interesting direction is to consider the incomplete data setting (missing at random, censored in different schemes) for locally spatial-functional data. A natural question is how to adapt our results to the wavelet-based estimators, the delta sequence estimators, the k NN estimators, and the local linear estimators.

9. Mathematical Developments

The proofs for our results are covered in this section. The following continues to use the notations that were previously presented.

To avoid the repetition of the Blocking technique and the notation used, we will devote the following subsection to introducing all notations needed for this decomposition.

9.1. Preliminaries

This treatment requires an extension of the Blocking techniques of Bernstein to the spacial process, refer to [85]. Let us introduce some notations related to this technique. Recall that $\{A_{1,n}\}$ and $\{A_{2,n}\}$ are sequences of positive numbers such that

$$A_{1,n}/A_n + A_{2,n}/A_{1,n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Let

$$A_{3,n} = A_{1,n} + A_{2,n}.$$

We consider a partition of \mathbb{R}^d by hypercubes of the form $\Gamma_n(\ell; \mathbf{0}) = (\ell + (0, 1]^d) A_{3,n}$, $\ell = (\ell_1, \dots, \ell_d)' \in \mathbb{Z}^d$ and divide $\Gamma_n(\ell; \mathbf{0})$ into 2^d hypercubes as follows:

$$\Gamma_n(\ell; \epsilon) = \prod_{j=1}^d I_j(\epsilon_j), \epsilon = (\epsilon_1, \dots, \epsilon_d)' \in \{1, 2\}^d, \tag{50}$$

where for $j = 1, \dots, d$,

$$I_j(\epsilon_j) = \begin{cases} (\ell_j A_{3,n}, \ell_j A_{3,n} + A_{1,n}] & \text{if } \epsilon_j = 1, \\ (\ell_j A_{3,n} + A_{1,n}, (\ell_j + 1) A_{3,n}] & \text{if } \epsilon_j = 2. \end{cases} \tag{51}$$

We note that

$$|\Gamma_n(\ell; \epsilon)| = A_{1,n}^{q(\epsilon)} A_{2,n}^{d-q(\epsilon)} \tag{52}$$

for any $\ell \in \mathbb{Z}^d$ and $\epsilon \in \{1, 2\}^d$, where

$$q(\epsilon) = [\{1 \leq j \leq d : \epsilon_j = 1\}].$$

Let $\epsilon_0 = (1, \dots, 1)'$. The partitions $\Gamma_n(\ell; \epsilon_0)$ correspond to “large blocks” and the partitions $\Gamma(\ell; \epsilon)$ for $\epsilon \neq \epsilon_0$ correspond to “small blocks”.

Let

$$L_{1,n} = \{ \ell \in \mathbb{Z}^d : \Gamma_n(\ell, 0) \subset R_n \}$$

be the index set of all hypercubes $\Gamma_n(\ell, 0)$ that are contained in R_n , and let

$$L_{2,n} = \{ \ell \in \mathbb{Z}^d : \Gamma_n(\ell, 0) \cap R_n \neq \emptyset, \Gamma_n(\ell, 0) \cap R_n^c \neq \emptyset \}$$

denote the boundary hypercubes index set. Define $L_n = L_{1,n} \cup L_{2,n}$.

9.2. Proof of Proposition 1

As we mentioned, our statistic is a weighted U -statistic that can be decomposed into a sum of U -statistics using the Hoeffding decomposition. We will treat this decomposition detailed in Section 3.1 to achieve the desired results. In the mentioned section, we have seen that

$$\hat{\psi}(\mathbf{u}, \mathbf{x}) - \mathbb{E}(\hat{\psi}(\mathbf{u}, \mathbf{x})) = \hat{\psi}_1(\mathbf{u}, \mathbf{x}) + \hat{\psi}_2(\mathbf{u}, \mathbf{x}),$$

where the linear term $\hat{\psi}_1(\mathbf{u}, \mathbf{x})$ and the remainder term $\hat{\psi}_2(\mathbf{u}, \mathbf{x})$ are well defined in (24) and (26), respectively. We aim to prove that the linear term leads the rate of convergence of this statistic while the remaining one converges to zero almost surely as $n \rightarrow \infty$. We will deal with the first term in the decomposition. For $B = [0, 1]$, $\alpha_n = \sqrt{\log n / nh^{md} \phi^m(h)}$ and $\tau_n = \rho_n n^{1/\zeta}$, where ζ is a positive constant given in Assumption 6 part (i), with $\rho_n = (\log n)^{\zeta_0}$ for some $\zeta_0 > 0$. Define

$$\tilde{H}_1^{(\ell)}(z) := \tilde{H}^{(\ell)}(z) \mathbf{1}_{\{|W_{s_i, A_n}| \leq \tau_n\}}, \tag{53}$$

$$\tilde{H}_2(z) := \tilde{H}^{(\ell)}(z) \mathbf{1}_{\{|W_{s_i, A_n}| > \tau_n\}}, \tag{54}$$

and

$$\hat{\psi}_1^{(1)}(\mathbf{u}, \mathbf{x}) - \theta(\mathbf{i}) = \frac{1}{n} \sum_{i=1}^n \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \xi_{i_1} \dots \xi_{i_{\ell-1}} \xi_i \xi_{i_\ell} \dots \xi_{i_{m-1}} \tilde{H}_1^{(\ell)}(z),$$

$$\hat{\psi}_1^{(2)}(\mathbf{u}, \mathbf{x}) - \theta(\mathbf{i}) = \frac{1}{n} \sum_{i=1}^n \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \xi_{i_1} \dots \xi_{i_{\ell-1}} \xi_i \xi_{i_\ell} \dots \xi_{i_{m-1}} \tilde{H}_2^{(\ell)}(z).$$

Clearly, we have

$$\begin{aligned} & \hat{\psi}_1(\mathbf{u}, \mathbf{x}) - \mathbb{E}\hat{\psi}_1(\mathbf{u}, \mathbf{x}) \\ &= \left[\hat{\psi}_1^{(1)}(\mathbf{u}, \mathbf{x}) - \mathbb{E}\hat{\psi}_1^{(1)}(\mathbf{u}, \mathbf{x}) \right] + \left[\hat{\psi}_1^{(2)}(\mathbf{u}, \mathbf{x}) - \mathbb{E}\hat{\psi}_1^{(2)}(\mathbf{u}, \mathbf{x}) \right]. \end{aligned} \tag{55}$$

To begin, it is plain to see that

$$\begin{aligned}
 & P_{\cdot|S} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \widehat{\psi}_1^{(2)}(u, x) - \theta(i) \right| > \alpha_n \right) \\
 &= P_{\cdot|S} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \widehat{\psi}_1^{(2)}(u, x) - \theta(i) \right| > \alpha_n \right) \\
 &\quad \cap \left\{ \sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \bigcup_{i=1}^n |W_{s_i, A_n}| > \tau_n \right\} \cup \left\{ \sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \left\{ \bigcup_{i=1}^n |W_{s_i, A_n}| > \tau_n \right\}^c \right\} \\
 &\leq P_{\cdot|S} \left\{ \sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \widehat{\psi}_1^{(2)}(u, x, \varphi) - \theta(i) \right| > \alpha_n \right. \\
 &\quad \left. \cap \left\{ \sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \bigcup_{i=1}^n |W_{s_i, A_n}| > \tau_n \right\} \right\} \\
 &\quad + P_{\cdot|S} \left\{ \sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \widehat{\psi}_2^{(1)}(u, x, \varphi) - \theta(i) \right| > \alpha_n \right. \\
 &\quad \left. \cap \left\{ \left\{ \sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \bigcup_{i=1}^n |W_{s_i, A_n}| > \tau_n \right\}^c \right\} \right\} \\
 &\leq P_{\cdot|S} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} |W_{s_i, A_n}| > \tau_n \text{ for some } i = 1, \dots, n \right) + P_{\cdot|S}(\emptyset) \\
 &\leq \tau_n^{-\zeta} \sum_{i=1}^n E_{\cdot|S} \left[\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} |W_{s_i, A_n}|^\zeta \right] \leq n \tau_n^{-\zeta} = \rho_n^{-\zeta} \rightarrow 0.
 \end{aligned}$$

We infer that

$$E_{\cdot|S} \left[\left| \widehat{\psi}_1^{(2)}(u, x) \right| \right] \leq \frac{1}{n} \sum_{i=1}^n \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \zeta_{i_1} \dots \zeta_{i_{\ell-1}} \zeta_i \zeta_{i_\ell} \dots \zeta_{i_{m-1}} E_{\cdot|S} \left(\left| \tilde{H}_2^{(\ell)}(z) \right| \right),$$

where

$$\begin{aligned}
 & E_{\cdot|S} \left(\left| \tilde{H}_2^{(\ell)}(z) \right| \right) \\
 &= E_{\cdot|S} \left[\left| \frac{1}{\phi(h)} K_2 \left(\frac{d(x_i, X_{s_i, A_n})}{h} \right) W_{s_i, A_n} \times \int W_{s_{(1, \dots, \ell-1, \ell, \dots, m-1)}, A_n} \prod_{\substack{j=1 \\ j \neq i}}^{m-1} \frac{1}{\phi(h)} \right. \right. \\
 &\quad \left. \left. K_2 \left(\frac{d(x_j, v_{s_j, A_n})}{h} \right) \mathbb{P}(dv_1, \dots, dv_{\ell-1}, dv_\ell, \dots, dv_{m-1}) \right| \right] \\
 &\leq E_{\cdot|S} \left[\left| \frac{1}{\phi(h)} \left(K_2 \left(\frac{d(x_i, X_{s_i, A_n})}{h} \right) + K_2 \left(\frac{d(x_i, X_{u_i}(s_i))}{h} \right) \right. \right. \right. \\
 &\quad \left. \left. - K_2 \left(\frac{d(x_i, X_{u_i}(s_i))}{h} \right) \right) W_{s_i, A_n} \mathbf{1}_{\{|W_{s_i, A_n}| > \tau_n\}} \right| \right] \\
 &\leq \frac{\tau_n^{-(\zeta-1)}}{\phi(h)} E_{\cdot|S} \left[\left| K_2 \left(\frac{d(x_i, X_{s_i, A_n})}{h} \right) - K_2 \left(\frac{d(x_i, X_{u_i}(s_i))}{h} \right) \right| |W_{s_i, A_n}|^\zeta \right. \\
 &\quad \left. + \left| K_2 \left(\frac{d(x_i, X_{u_i}(s_i))}{h} \right) \right| |W_{s_i, A_n}|^\zeta \right] \tag{56}
 \end{aligned}$$

$$\begin{aligned} &\leq \frac{\tau_n^{-(\zeta-1)}}{\phi(h)} E_{\cdot|S} \left[h^{-1} |d(x_i, X_{s_i, A_n}) - d(x_i, X_{u_i}(s_i))| |W_{s_i, A_n}|^\zeta \right] \\ &\quad + E_{\cdot|S} \left[\left| K_2 \left(\frac{d(x_i, X_{u_i}(s_i))}{h} \right) \right| |W_{s_i, A_n}|^\zeta \right] \\ &\lesssim \frac{\tau_n^{-(\zeta-1)}}{\phi(h)} \times \left[\frac{1}{nh} + \phi(h) \right] \\ &\lesssim \frac{\tau_n^{-(\zeta-1)}}{nh\phi(h)} + \tau_n^{-(\zeta-1)}. \end{aligned}$$

Hence, we have

$$\begin{aligned} E_{\cdot|S} \left[\left| \widehat{\psi}_1^{(2)}(\mathbf{u}, \mathbf{x}) \right| \right] &\lesssim \frac{1}{n} \sum_{i=1}^n \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \xi_{i_1} \cdots \xi_{i_{\ell-1}} \xi_i \xi_{i_\ell} \cdots \xi_{i_{m-1}} \left[\tau_n^{-(\zeta-1)} \right] \\ &\lesssim \tau_n^{-(\zeta-1)} \frac{1}{n^m} \sum_{i \in I_n^m} \prod_{j=1}^m \frac{1}{h^d} \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_j / A_n}{h_n} \right) \\ &= \frac{C}{\tau_n^{(\zeta-1)}} \left(f_S(\mathbf{u}) + O \left(\sqrt{\frac{\log n}{nh^{md}}} + h^2 \right) \right) \text{ (Using Lemma A1)} \\ &\leq \frac{C}{\tau_n^{(\zeta-1)}} = C\rho_n^{-(\zeta-1)} n^{-(\zeta-1)/\zeta} \leq C\alpha_n \quad P_S - a.s. \end{aligned}$$

As a result, we obtain that

$$\sup_{\mathcal{F}_m} \sup_{\mathcal{H}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \left| \widehat{\psi}_1^{(2)}(\mathbf{u}, \mathbf{x}) - E_{\cdot|S} \widehat{\psi}_1^{(2)}(\mathbf{u}, \mathbf{x}) \right| = O_{P_{\cdot|S}}(\alpha_n). \tag{57}$$

Second, let us treat

$$\sup_{\mathcal{F}_m} \sup_{\mathcal{H}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \left| \widehat{\psi}_1^{(1)}(\mathbf{u}, \mathbf{x}, \varphi) - \mathbb{E} \widehat{\psi}_1^{(1)}(\mathbf{u}, \mathbf{x}, \varphi) \right|.$$

Recall the large blocks and small blocks and the notation given in Section 9.1, and define

$$\begin{aligned} S_{s, A_n}(\mathbf{u}, \mathbf{x}) &:= \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \xi_{i_1} \cdots \xi_{i_{\ell-1}} \xi_i \xi_{i_\ell} \cdots \xi_{i_{m-1}} \tilde{H}_1^{(\ell)}(z), \\ S_n(\boldsymbol{\ell}; \boldsymbol{\epsilon}) &= \sum_{i: s_i \in \Gamma_n(\boldsymbol{\ell}; \boldsymbol{\epsilon}) \cap R_n} S_{s, A_n}(\mathbf{u}, \mathbf{x}) = \left(S_n^{(1)}(\boldsymbol{\ell}; \boldsymbol{\epsilon}), \dots, S_n^{(p)}(\boldsymbol{\ell}; \boldsymbol{\epsilon}) \right)'. \end{aligned}$$

Then, we have

$$\begin{aligned} S_n &= \left(S_n^{(1)}, \dots, S_n^{(m)} \right)' \\ &= \sum_{i=1}^n S_{s, A_n}(\mathbf{u}, \mathbf{x}) \\ &= \sum_{\boldsymbol{\ell} \in L_n} S_n(\boldsymbol{\ell}; \boldsymbol{\epsilon}_0) + \underbrace{\sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} \sum_{\boldsymbol{\ell} \in L_{1,n}} S_n(\boldsymbol{\ell}; \boldsymbol{\epsilon})}_{=: S_{2,n}(\boldsymbol{\epsilon})} + \underbrace{\sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} \sum_{\boldsymbol{\ell} \in L_{2,n}} S_n(\boldsymbol{\ell}; \boldsymbol{\epsilon})}_{=: S_{3,n}(\boldsymbol{\epsilon})} \tag{58} \\ &=: S_{1,n} + \sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} S_{2,n}(\boldsymbol{\epsilon}) + \sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} S_{3,n}(\boldsymbol{\epsilon}). \end{aligned}$$

In order to achieve our result, we will pass by the following two steps.

Step 1 (Reduction to independence). Recall

$$S_n(\ell; \epsilon) = \sum_{i: s_i \in \Gamma_n(\ell; \epsilon) \cap R_n} S_{s_i, A_n}(\mathbf{u}, \mathbf{x}).$$

For each $\epsilon \in \{1, 2\}^d$, let $\{\check{S}_n(\ell; \epsilon) : \ell \in L_n\}$ be a sequence of independent random variables in \mathbb{R} under $P_{\cdot|S}$ such that

$$\check{S}_n(\ell; \epsilon) \stackrel{d}{=} S_n(\ell; \epsilon), \text{ under } P_{\cdot|S}, \ell \in L_n.$$

Define

$$\check{S}_{1,n} = \sum_{\ell \in L_n} \check{S}_n(\ell; \epsilon_0) = \left(\check{S}_{1,n}^{(1)}, \dots, \check{S}_{1,n}^{(m)} \right)'$$

and for $\epsilon \neq \epsilon_0$, define

$$\check{S}_{2,n}(\epsilon) = \sum_{\ell \in L_{1,n}} \check{S}_n(\ell; \epsilon)$$

and

$$\check{S}_{3,n}(\epsilon) = \sum_{\ell \in L_{2,n}} \check{S}_n(\ell; \epsilon).$$

We start by confirming the following results:

$$\sup_{t>0} \left| P_{\cdot|S}(S_{1,n} > t) - P_{\cdot|S}(\check{S}_{1,n} > t) \right| \leq C \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d), \tag{59}$$

$$\sup_{t>0} \left| P_{\cdot|S}(\|S_{2,n}(\epsilon)\|_\infty > t) - P_{\cdot|S}(\|\check{S}_{2,n}(\epsilon)\|_\infty > t) \right| \leq C \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d), \tag{60}$$

$$\sup_{t>0} \left| P_{\cdot|S}(\|S_{3,n}(\epsilon)\|_\infty > t) - P_{\cdot|S}(\|\check{S}_{3,n}(\epsilon)\|_\infty > t) \right| \leq C \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d). \tag{61}$$

Keep in mind that

$$\llbracket L_n \rrbracket = O\left((A_n / A_{3,n})^d \right) \lesssim (A_n / A_{1,n})^d.$$

For $\epsilon \in \{1, 2\}^d$ and $\ell_1, \ell_2 \in L_n$ with $\ell_1 \neq \ell_2$, let

$$\mathcal{J}_1(\epsilon) = \{1 \leq i_1 \leq n : s_{i_1} \in \Gamma_n(\ell_1; \epsilon)\}, \mathcal{J}_2(\epsilon) = \{1 \leq i_2 \leq n : s_{i_2} \in \Gamma_n(\ell_2; \epsilon)\}.$$

For any $s_{i_k} = (s_{1,i_k}, \dots, s_{d,i_k}), k = 1, 2$ in such a way that $i_1 \in \mathcal{J}_1(\epsilon)$ and $i_2 \in \mathcal{J}_2(\epsilon)$, we obtain $\max_{1 \leq u \leq d} |s_{u,i_1} - s_{u,i_2}| \geq A_{2,n}$ using the definition of $\Gamma(\ell; \epsilon)$. This gives

$$|s_{i_1} - s_{i_2}| \geq A_{2,n}.$$

For any $\epsilon \in \{1, 2\}^d$, let $S_n(\ell_1; \epsilon), \dots, S_n(\ell_{\llbracket L_n \rrbracket}; \epsilon)$ be an arrangement of $\{S_n(\ell; \epsilon) : \ell \in L_n\}$.

Let $P_{\cdot|S}^{(a)}$ be the marginal distribution of $S_n(\ell_a; \epsilon)$ and let $P_{\cdot|S}^{(a:b)}$ be the joint distribution of $\{S_n(\ell_k; \epsilon) : a \leq k \leq b\}$. The β -mixing property of X gives that for $1 \leq k \leq \llbracket L_n \rrbracket - 1$,

$$\left\| P_{\cdot|S} - P_{\cdot|S}^{(1:k)} \times P_{\cdot|S}^{(k+1:\llbracket L_n \rrbracket)} \right\|_{TV} \lesssim \beta(A_{2,n}; A_n^d).$$

The inequality is independent of the arrangement of $\{S_n(\ell; \epsilon) : \ell \in L_n\}$. Therefore, the Assumption A2 in Lemma A4 is fulfilled for $\{S_n(\ell; \epsilon) : \ell \in L_n\}$ with $\tau \sim \beta(A_{2,n}; A_n^d)$ and $m \lesssim (A_n / A_{1,n})^d$. Combining the boundary condition on R_n and Lemma A4, we get (59)–(61).

Remark 10. Since

$$\llbracket \{ \epsilon \in \{1, 2\}^d : \epsilon \neq \epsilon_0 \} \rrbracket = 2^d - 1, \llbracket L_{1,n} \rrbracket \sim (A_n / A_{3,n})^d \sim (A_n / A_{1,n})^d$$

and

$$\llbracket L_{2,n} \rrbracket \sim (A_n / A_{3,n})^{d-1} \sim (A_n / A_{1,n})^{d-1} \ll \llbracket L_{1,n} \rrbracket,$$

Lemma A5 and Equation (52) give for sufficiently large n the summands numbers of $S_{2,n}$ and $S_{3,n}$ are at most

$$O\left(A_{1,n}^{d-1} A_{2,n} n A_n^{-d} (A_n / A_{1,n})^d\right) = O\left(\frac{A_{2,n}}{A_{1,n}} n\right)$$

and

$$O\left(A_{1,n}^{d-1} A_{2,n} n A_n^{-d} (A_n / A_{1,n})^{d-1}\right) = O\left(\frac{A_{2,n}}{A_n} n\right),$$

respectively.

Step 2. Recall that we aim to treat

$$\sup_{\mathcal{F}_m} \sup_{\mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \widehat{\psi}_1^{(1)}(u, x, \varphi) - E_{\cdot|S} \widehat{\psi}_1^{(1)}(u, x, \varphi) \right|.$$

To achieve the intended result, we will cover the region $B^m = [0, 1]^{dm}$ by

$$\bigcup_{k_1, \dots, k_m=1}^{N(u)} \prod_{j=1}^m B(u_{k_j}, r),$$

for some radius r . Hence, for each $u = (u_1, \dots, u_m) \in [0, 1]^{dm}$, there exists $l(u) = (l(u_1), \dots, l(u_m))$, where $\forall 1 \leq i \leq m, 1 \leq l(u_i) \leq N(u_i)$ in such a way that

$$u \in \prod_{i=1}^m B(u_{l(u_i)}, r) \text{ and } |u_i - u_{l(u_i)}| \leq r, \text{ for } 1 \leq i \leq m,$$

then for each $u \in [0, 1]^{dm}$, the closest center will be $u_l(u)$, and the ball with the closest center will be defined by

$$\mathcal{B}(u, l(u), r) := \prod_{j=1}^m B(u_{k_j}, r).$$

In the same way, \mathcal{H}^m should be covered by

$$\bigcup_{k_1, \dots, k_m=1}^{N(x)} \prod_{j=1}^m B(x_{k_j}, r),$$

for some radius r . Hence, for each $x = (x_1, \dots, x_m) \in \mathcal{H}^m$, there exists $l(x) = (l(x_1), \dots, l(x_m))$, where $\forall 1 \leq i \leq m, 1 \leq l(x_i) \leq N(x_i)$ in such a way that

$$x \in \prod_{i=1}^m B(u_{l(x_i)}, r) \text{ and } d(x_i, x_{l(x_i)}) \leq r, \text{ for } 1 \leq i \leq m,$$

then for each $x \in \mathcal{H}^m$, the closest center will be $x_l(x)$, and the ball with the closest center will be defined by

$$B(x, l(x), r) := \prod_{i=1}^m B(x_{l(x_i)}, r).$$

We define:

$$K^*(\omega, v) \geq C_0 \prod_{j=1}^m \prod_{\ell=1}^d \mathbf{1}_{(|\omega_{j,\ell}| \leq 2C_1)} \prod_{j=1}^m K_2(v_k) \text{ for } (\omega, v) \in \mathbb{R}^2.$$

We can show that for $(u, x) \in B_{j,n}$ and n large enough,

$$\left| \bar{K}\left(\frac{u-s/A_n}{h_n}\right) K_2\left(\frac{d(x_i, X_{s_i, A_n})}{h}\right) - \bar{K}\left(\frac{u_n-s/A_n}{h_n}\right) K_2\left(\frac{d(x_{n,i}, X_{s_i, A_n})}{h}\right) \right| \leq \alpha_n K^*\left(\frac{u_n-s/A_n, d(x_{n,i}, X_{s_i, A_n})}{h_n}\right).$$

Then, for

$$\begin{aligned} \hat{\psi}_1^{(1)}(u, x) &= \frac{1}{n} \sum_{i=1}^n \zeta_i \frac{1}{\phi(h)} K_2\left(\frac{d(x_i, X_{s_i, A_n})}{h}\right) W_{s_i, A_n} \mathbf{1}_{\{|W_{s_i, A_n}| \leq \tau_n\}} \\ &\times \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \zeta_{i_1} \dots \zeta_{i_{\ell-1}} \zeta_{i_\ell} \dots \zeta_{i_{m-1}} \\ &\times \int W_{s_{(1, \dots, \ell-1, \ell, \dots, m-1)}, A_n} \prod_{\substack{j=1 \\ j \neq i}}^{m-1} \frac{1}{\phi(h)} K_2\left(\frac{d(x_j, v_{s_j, A_n})}{h}\right) \mathbb{P}(dv_1, \dots, dv_{\ell-1}, dv_\ell, \dots, dv_{m-1}). \end{aligned}$$

Let us define

$$\begin{aligned} \bar{\psi}_1^{(1)}(u, x) &= \frac{1}{nh^d \phi(h)} \sum_{i=1}^n K^*\left(\frac{u_n-s_i/A_n, d(x_{n,i}, X_{s_i, A_n})}{h_n}\right) W_{s_i, A_n} \mathbf{1}_{\{|W_{s_i, A_n}| \leq \tau_n\}} \\ &\times \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \zeta_{i_1} \dots \zeta_{i_{\ell-1}} \zeta_{i_\ell} \dots \zeta_{i_{m-1}} \\ &\times \int W_{s_{(1, \dots, \ell-1, \ell, \dots, m-1)}, A_n} \prod_{\substack{j=1 \\ j \neq i}}^{m-1} \frac{1}{\phi(h)} K_2\left(\frac{d(x_j, v_{s_j, A_n})}{h}\right) \mathbb{P}(dv_1, \dots, dv_{\ell-1}, dv_\ell, \dots, dv_{m-1}) \\ &:= \frac{1}{nh^d \phi(h)} \sum_{i=1}^n S'_{s_i, A_n}(u, x). \end{aligned} \tag{62}$$

We mention that

$$E_{\cdot|S} \left[\left| \bar{\psi}_1^{(1)}(u, x, \varphi) \right| \right] \leq M < \infty,$$

for some M large enough. Let $N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m$ denote the covering number related, respectively, to the class of functions $\mathcal{F}_m \mathcal{K}^m$, the balls that cover $[0, 1]^m$ and the balls that cover \mathcal{H}^m . Then, we obtain

$$\begin{aligned} &\sup_{\mathcal{F}_m \mathcal{K}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in \mathcal{B}} \left| \hat{\psi}_1^{(1)}(u, x, \varphi) - E_{\cdot|S} \left[\hat{\psi}_1^{(1)}(u, x, \varphi) \right] \right| \\ &\leq N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x), r})} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u), r})} \left| \hat{\psi}_1^{(1)}(u, x, \varphi) - E_{\cdot|S} \left[\hat{\psi}_1^{(1)}(u, x, \varphi) \right] \right| \\ &\leq N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x), r})} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u), r})} \left| \hat{\psi}_1^{(1)}(u_n, x) - E_{\cdot|S} \left[\hat{\psi}_1^{(1)}(u_n, x) \right] \right| \\ &\quad + N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x), r})} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u), r})} \alpha_n \left(\left| \bar{\psi}_1^{(1)}(u_n, x) \right| + E_{\cdot|S} \left[\left| \bar{\psi}_1^{(1)}(u_n, x) \right| \right] \right) \\ &\leq N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x), r})} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u), r})} \left| \hat{\psi}_1^{(1)}(u_n, x) - E_{\cdot|S} \left[\hat{\psi}_1^{(1)}(u_n, x) \right] \right| \end{aligned}$$

$$\begin{aligned}
 & + N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \bar{\psi}_1^{(1)}(\mathbf{u}_n, \mathbf{x}) - E_{\cdot|\mathcal{S}} \left[\bar{\psi}_1^{(1)}(\mathbf{u}_n, \mathbf{x}) \right] \right| \\
 & + 2MF(\mathbf{y})\alpha_n \\
 \leq & N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \sum_{\ell \in L_n} S_n(\ell; \epsilon_0) \right| \tag{63} \\
 & + N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \sum_{\epsilon \neq \epsilon_0} \sum_{\ell \in L_{1,n}} S_n(\ell; \epsilon) \right| \\
 & + N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \sum_{\epsilon \neq \epsilon_0} \sum_{\ell \in L_{2,n}} S_n(\ell; \epsilon) \right| \\
 & + N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \sum_{\ell \in L_n} S'_n(\ell; \epsilon_0) \right| \\
 & + N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \sum_{\epsilon \neq \epsilon_0} \sum_{\ell \in L_{1,n}} S'_n(\ell; \epsilon) \right| \\
 & + N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \sum_{\epsilon \neq \epsilon_0} \sum_{\ell \in L_{2,n}} S'_n(\ell; \epsilon) \right| + 2MF(\mathbf{y})\alpha_n.
 \end{aligned}$$

Even more, for each $\epsilon \in \{1, 2\}^d$, let $\{\check{S}'_n(\ell; \epsilon) : \ell \in L_n\}$ denote a sequence of independent random vectors in \mathbb{R}^m under $P_{\cdot|\mathcal{S}}$ such that

$$\check{S}'_n(\ell; \epsilon) \stackrel{d}{=} S'_n(\ell; \epsilon), \text{ under } P_{\cdot|\mathcal{S}}, \ell \in L_n.$$

Show that

$$\begin{aligned}
 & P_{\cdot|\mathcal{S}} \left(N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \right. \\
 & \quad \left. \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \left| \widehat{\psi}_1^{(1)}(\mathbf{u}, \mathbf{x}) - E_{\cdot|\mathcal{S}} \left[\widehat{\psi}_1^{(1)}(\mathbf{u}, \mathbf{x}) \right] \right| > 2^{md+1} Ma_n \right) \tag{64}
 \end{aligned}$$

$$\begin{aligned}
 \leq & N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} \\
 & P_{\cdot|\mathcal{S}} \left(\sup_{(\mathbf{u}, \mathbf{x}) \in B_k} \left| \widehat{\psi}_1(\mathbf{u}, \mathbf{x}) - E_{\cdot|\mathcal{S}} \left[\widehat{\psi}_1(\mathbf{u}, \mathbf{x}) \right] \right| > 2^{md+1} Ma_n \right) \\
 \leq & \sum_{\epsilon \in \{1, 2\}^d} \widehat{Q}_n(\epsilon) + \sum_{\epsilon \in \{1, 2\}^d} \bar{Q}_n(\epsilon) + 2^{md+1} N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d), \tag{65}
 \end{aligned}$$

where

$$\begin{aligned}
 \widehat{Q}_n(\epsilon_0) &= N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \\
 & \quad \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} P_{\cdot|\mathcal{S}} \left(\left| \sum_{\ell \in L_n} \check{S}_n(\ell; \epsilon_0) \right| > Ma_n n^m h^{md} \phi(h) \right), \\
 \bar{Q}_n(\epsilon_0) &= N_{\mathcal{F}_m \mathcal{K}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)}
 \end{aligned}$$

$$\max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} P_{\cdot|S} \left(\left| \sum_{\ell \in L_n} \check{S}'_n(\ell; \epsilon_0) \right| > Ma_n n^m h^{md} \phi(h) \right),$$

and for $\epsilon \neq \epsilon_0$

$$\widehat{Q}_n(\epsilon) = N_{\mathcal{F}_m \mathcal{X}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} P_{\cdot|S} \left(\left| \sum_{\ell \in L_n} \check{S}_n(\ell; \epsilon) \right| > Ma_n n^m h^{md} \phi(h) \right),$$

$$\widehat{Q}_n(\epsilon) = N_{\mathcal{F}_m \mathcal{X}^m} N_{(x)}^m N_{(u)}^m \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(x_{i(x)}, r)} \max_{1 \leq i_1 < \dots < i_m \leq m} \sup_{B(u_{i(u)}, r)} P_{\cdot|S} \left(\left| \sum_{\ell \in L_n} \check{S}'_n(\ell; \epsilon) \right| > Ma_n n^m h^{md} \phi(h) \right).$$

Due to the similarity between the two cases, $\epsilon \neq \epsilon_0$ and $\epsilon = \epsilon_0$, we are going to treat \widehat{Q}_n only for $\epsilon \neq \epsilon_0$. An application of Lemma A5, with the fact that $\check{S}_n(\ell; \epsilon)$ are zero-mean random variables, shows us that:

$$P_{\cdot|S} \left(\left| \sum_{\ell \in L_n} \check{S}_n(\ell; \epsilon) \right| > Ma_n n h^{md} \phi(h) \right) \leq 2P_{\cdot|S} \left(\sum_{\ell \in L_n} \check{S}_n(\ell; \epsilon) > Ma_n n h^{md} \phi(h) \right)$$

and

$$\begin{aligned} |\check{S}_n(\ell; \epsilon)| &\leq CA_{1,n}^{d-1} A_{2,n} (\log n) \tau_n, P_S - \text{ a.s. (from Lemma A5)} \\ E_{\cdot|S} \left[(\check{S}_n(\ell; \epsilon))^2 \right] &\leq Ch^{md} \phi(h) A_{1,n}^{d-1} A_{2,n} (\log n), P_S - \text{ a.s. (By Lemma A6)} \end{aligned} \tag{66}$$

Using Bernstein’s inequality represented in Lemma A7, we have

$$\begin{aligned} &P_{\cdot|S} \left(\sum_{\ell \in L_n} \check{S}_n(\ell; \epsilon) > Ma_n n h^{md} \phi(h) \right) \\ &\leq \exp \left(- \frac{\frac{1}{2} \times Mn h^{md} \phi(h) \log n}{\left(\frac{A_n}{A_{1,n}} \right)^d A_{1,n}^{d-1} A_{2,n} h^{md} \phi(h) (\log n) + \frac{1}{3} \times M^{1/2} n^{1/2} h^{md/2} \phi(h)^{1/2} (\log n)^{1/2} A_{1,n}^{d-1} A_{2,n} \tau_n} \right). \end{aligned}$$

Observe that

$$\frac{nh^{md} \log n}{\left(\frac{A_n}{A_{1,n}} \right)^d A_{1,n}^{d-1} A_{2,n} h^{md} (\log n)} = n A_n^{-d} \left(\frac{A_{1,n}}{A_{2,n}} \right) \gtrsim \frac{A_{1,n}}{A_{2,n}} \gtrsim n^\eta, \tag{67}$$

$$\frac{nh^{md} \phi(h) \log n}{n^{1/2} h^{md/2} \phi(h)^{1/2} (\log n)^{1/2} A_{1,n}^{d-1} A_{2,n} \tau_n} = \frac{n^{1/2} h^{md/2} \phi(h)^{1/2} (\log n)^{1/2}}{A_{1,n}^{md} \left(\frac{A_{2,n}}{A_{1,n}} \right) \rho_n n^{1/\zeta}} \geq C_0 n^{\eta/2}. \tag{68}$$

Taking $M > 0$ to be sufficiently large, and for $N \leq Ch^{-md} \phi(h) \alpha_n^{-m}$, this shows the desired result.

We must move on to the nonlinear part of the Hoeffding decomposition. Accordingly, the goal is to prove that:

$$P_{\cdot|S} \left[\sup_{\mathcal{F}_m \mathcal{K}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} |\widehat{\psi}_2(\mathbf{u}, \mathbf{x})| > \lambda \right] \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{69}$$

In the following, we will give a lemma that can be viewed as a technical result in the proof of our proposition, and it helps us to achieve our goal in Expression (69). The proof of this lemma used the Blocking technique defined before but for the U -statistic, making the block treatment more complicated.

Lemma 1. *Let $\mathcal{F}_m \mathcal{K}^m$ be a uniformly bounded class of measurable canonical functions, $m \geq 2$. Assume that there are finite constants \mathbf{a} and \mathbf{b} in such a way that the $\mathcal{F}_m \mathcal{K}^m$ covering number fulfills:*

$$N(\epsilon, \mathcal{F}_m \mathcal{K}^m, \|\cdot\|_{L_2(Q)}) \leq \mathbf{a}\epsilon^{-\mathbf{b}}, \tag{70}$$

for all $\epsilon > 0$ and all probability measures Q . If the mixing coefficients β of the local stationary sequence $\{Z_i = (X_{s_i, A_n}, W_{s_i, A_n})\}_{i \in \mathbb{N}^*}$ satisfy Condition (E2) in Assumption 6, then, for some $r > 1$, we have:

$$\sup_{\mathcal{F}_m \mathcal{K}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \mathbb{P} \left[h^{md/2} \phi^{m/2}(h) n^{-m+1/2} \sum_{i \in I_n^m} \xi_{i_1} \dots \xi_{i_m} H(Z_{i_1}, \dots, Z_{i_m}) \right] \rightarrow 0. \tag{71}$$

Remark 11. *As mentioned before, W_{s_i, A_n} will be equal to 1 or $\epsilon_{s_i, A_n} = \sigma\left(\frac{s_i}{A_n}, X_{s_i, A_n}\right) \epsilon_{s_i}$. In the proof of the previous lemma, W_{s_i, A_n} will be equal $\epsilon_{i, n} = \sigma\left(\frac{i}{n}, X_{i, n}\right) \epsilon_i$, and we will use the notation $W_{s_i, A_n}^{(u)}$ to indicate $\sigma(\mathbf{u}, \mathbf{x}) \epsilon_i$.*

□

9.2.1. Proof of Lemma 1

This lemma’s proof is based on the blocking technique employed by [82], and it is called *Bernstein’s method*, referred to [150], in which we are enabled to apply the symmetrization and the many other techniques available for the i.i.d random variables. We will extend this technique to the spacial processes in the U -statistics setting, in the same line as in [93]. In addition to the notation in Section 9.1, define

$$L_n := L_{1,n} \cup L_{2,n},$$

$$\Delta_1 = \{\ell_2 : \min_{1 \leq i \leq d} |\ell_{1i} - \ell_{2i}| \leq 1\}$$

$$\Delta_2 = \{\ell_2 : \min_{1 \leq i \leq d} |\ell_{1i} - \ell_{2i}| \geq 2\}$$

With the notation introduced above, it is easy to show that, for $m = 2$,

$$\begin{aligned} & \frac{1}{h^{2d} \phi^2(h)} \sum_{i \in I_n^2} \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\} W_{s_i, A_n} \\ &= \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \\ & \quad \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\} W_{s_i, A_n} \\ &+ \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \end{aligned}$$

$$\begin{aligned}
 & \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_{i_j}, A_n} \\
 +2 & \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: \mathbf{s}_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: \mathbf{s}_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \\
 & \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_{i_j}, A_n} \\
 +2 & \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: \mathbf{s}_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_1} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: \mathbf{s}_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \\
 & \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_{i_j}, A_n} \\
 + & \frac{1}{h^{2d} \phi^2(h)} \sum_{\substack{\ell_1 \neq \ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_1: \mathbf{s}_{i_1} \in \Gamma_n(\ell_1; \epsilon) \cap R_n} \sum_{i_2: \mathbf{s}_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \\
 & \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_{i_j}, A_n} \\
 + & \frac{1}{h^{2d} \phi^2(h)} \sum_{\substack{\ell_1 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_1 < i_2: \mathbf{s}_{i_1}, \mathbf{s}_{i_2} \in \Gamma_n(\ell_1; \epsilon) \cap R_n} \\
 & \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_{i_j}, A_n} \\
 := & \text{I} + \text{II} + \text{III} + \text{IV} + \text{V} + \text{VI}.
 \end{aligned} \tag{72}$$

(I): The Same Type of Blocks but Not the Same Block

Let $\{\eta_i\}_{i \in \mathbb{N}^*}$ be a sequence of independent blocks. An application of Lemma A4 shows that:

$$\begin{aligned}
 & \mathbb{P} \left(\sup_{\mathcal{F}_m} \sup_{\mathcal{X}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \left| n^{-3/2} \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: \mathbf{s}_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: \mathbf{s}_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \right. \right. \\
 & \quad \left. \left. \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_{i_j}, A_n} \right| > \delta \right) \\
 & \leq \mathbb{P} \left(\sup_{\mathcal{F}_m} \sup_{\mathcal{X}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: \mathbf{s}_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: \mathbf{s}_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \right. \right. \\
 & \quad \left. \left. \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{u}_j}(\mathbf{s}_{i_j}))}{h} \right) \right] W_{\mathbf{s}_{i_j}, A_n} \right| > \delta \right) \\
 + & \mathbb{P} \left(\sup_{\mathcal{F}_m} \sup_{\mathcal{X}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: \mathbf{s}_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: \mathbf{s}_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \right. \right. \\
 & \quad \left. \left. \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{u}_j}(\mathbf{s}_{i_j}))}{h} \right) \left[W_{\mathbf{s}_{i_j}, A_n} - W_{\mathbf{s}_{i_j}, A_n}^{(u)} \right] \right| > \delta \right)
 \end{aligned}$$

$$\begin{aligned}
 & + \mathbb{P} \left(\sup_{\mathcal{F}_m} \sup_{\mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in \mathcal{B}^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \right. \right. \\
 & \qquad \left. \left. \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_i, X_{u_j}(s_{i_j}))}{h} \right) W_{s_i, A_n}^{(u)} \right| > \delta \right) \\
 & \leq \mathbb{P} \left(\sup_{\mathcal{F}_m} \sup_{\mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in \mathcal{B}^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \right. \right. \\
 & \qquad \left. \left. \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right| > \delta \right) \\
 & \quad + C \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d) + o_{\mathbb{P}}(1) + o_{\mathbb{P}}(1),
 \end{aligned}$$

Because:

$$\begin{aligned}
 & E_{\cdot|S} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \right. \\
 & \qquad \left. \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) \right] W_{s_i, A_n} \right| \\
 & = \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \\
 & \qquad E_{\cdot|S} \left| \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) \right] W_{s_i, A_n} \right| \\
 & = \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \\
 & \qquad E_{\cdot|S} \left| \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) \right] \prod_{j=1}^m \epsilon_{s_{i_j}, A_n} \right| \\
 & = \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \\
 & \qquad E_{\cdot|S} \left| \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) \right] \prod_{j=1}^m \sigma \left(\frac{s_{i_j}}{A_n}, X_{s_{i_j}, A_n} \right) \epsilon_{s_{i_j}} \right| \\
 & = \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^m E_{\cdot|S}(\epsilon_{s_{i_j}}) \tag{73} \\
 & \qquad E_{\cdot|S} \left| \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) \right] \right. \\
 & \qquad \left. \left[\prod_{j=1}^m \sigma \left(\frac{s_{i_j}}{A_n}, X_{s_{i_j}, A_n} \right) - \prod_{j=1}^m \sigma(u_j, x_j) + \prod_{j=1}^m \sigma(u_j, x_j) \right] \right|
 \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^m E_{\cdot|S}(\epsilon_{s_{i_j}}) \\ &\quad E_{\cdot|S} \left[\left| C \sum_{j=1}^m K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - K_2 \left(\frac{d\left(x_j, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h} \right) \right|^p \right] \prod_{j=1}^m [\sigma(\mathbf{u}_j, \mathbf{x}_j) + o_{\mathbb{P}}(1)] \\ &\quad \text{(Using a telescoping argument, and the boundedness of } K_2 \text{ for } p = \min(\rho, 1) \text{ and } C < \infty) \\ &\lesssim \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^m E_{\cdot|S}(\epsilon_{s_{i_j}}) \\ &\quad E_{\cdot|S} \left[\phi^{m-1}(h) \left| \frac{C}{A_n^d} U_{s_{i_j}, A_n} \left(\frac{\mathbf{s}_{i_j}}{A_n} \right) \right|^p \right] \prod_{j=1}^m [\sigma(\mathbf{u}_j, \mathbf{x}_j) + o_{\mathbb{P}}(1)] \\ &\sim o_{\mathbb{P}}(1), \end{aligned}$$

and

$$\begin{aligned} &E_{\cdot|S} \left[\frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \right. \\ &\quad \left. \prod_{j=1}^2 K_2 \left(\frac{d(x_i, X_{u_j}(s_{i_j}))}{h} \right) \left[W_{s_{i_j}, A_n} - W_{s_{i_j}, A_n}^{(u)} \right] \right] \\ &= \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^m E_{\cdot|S}(\epsilon_{s_{i_j}}) \\ &\quad E_{\cdot|S} \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_i, X_{u_j}(s_{i_j}))}{h} \right) \left[\prod_{j=1}^m \sigma\left(\frac{\mathbf{s}_{i_j}}{A_n}, X_{s_{i_j}, A_n}\right) - \prod_{j=1}^m \sigma(\mathbf{u}_j, \mathbf{x}_j) \right] \right] \\ &\lesssim \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^m E_{\cdot|S}(\epsilon_{s_{i_j}}) \tag{74} \\ &\quad \times (o_{\mathbb{P}}(1)) \int_0^h \prod_{k=1}^m K_2\left(\frac{y_k}{h}\right) dF_{i_k/n}(y_k, x_k) \\ &\lesssim \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^m E_{\cdot|S}(\epsilon_{s_{i_j}}) \\ &\quad \times (o_{\mathbb{P}}(1)) (\phi^2(h)) \\ &\sim o_{\mathbb{P}}(1). \end{aligned}$$

Under the assumptions of the lemma, we have $\beta(a; b) \leq \beta_1(a)g_1(b)$ with $\beta_1(a) \rightarrow 0$ as $a \rightarrow \infty$ and $n \rightarrow \infty$, so the term to consider is the first summand. For the second part of the inequality, we will use the work of [27] in the non-fixed kernels settings, precisely, we will define $f_{i_1, \dots, i_m} = \prod_{k=1}^m \zeta_{i_k} \times H$ and $\mathcal{F}_{i_1, \dots, i_m}$, respectively, as a collection of kernels and the class of functions related to this kernel; then, we will use ([29], Theorem 3.1.1 and Remarks 3.5.4 part 2) for decoupling and randomization. As we mentioned above, we will suppose that $m = 2$. Then, we can see that

$$E_{\cdot|S} \left\| \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \right\|$$

$$\begin{aligned}
 & \left\| \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{i, \varphi, n}^{(u)} \right\|_{\mathcal{F}_2 \mathcal{K}^2} \\
 = & E_{\cdot | S} \left\| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} f_{i_1, i_2}(\mathbf{u}, \boldsymbol{\eta}) \right\|_{\mathcal{F}_{i_1, i_2}} \tag{75} \\
 \leq & c_2 E_{\cdot | S} \left\| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \epsilon_p \epsilon_q \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} f_{i_1, i_2}(\mathbf{u}, \boldsymbol{\eta}) \right\|_{\mathcal{F}_{i_1, i_2}} \\
 \leq & c_2 E_{\cdot | S} \int_0^{D_{nh}^{(U_1)}} N(t, \mathcal{F}_{i_1, i_2}, \tilde{d}_{nh, 2}^{(1)}) dt, \quad (\text{By Lemma A9 and Proposition A1.})
 \end{aligned}$$

where $D_{nh}^{(U_1)}$ is the diameter of \mathcal{F}_{i_1, i_2} according to the distance $\tilde{d}_{nh, 2}^{(1)}$, respectively, which is defined as

$$\begin{aligned}
 D_{nh}^{(U_1)} & := \left\| \mathbb{E}_\epsilon \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \epsilon_p \epsilon_q \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} f_{i_1, i_2}(\mathbf{u}, \boldsymbol{\eta}) \right| \right\|_{\mathcal{F}_{i_1, i_2}} \\
 & = \left\| \mathbb{E}_\epsilon \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \epsilon_p \epsilon_q \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - s_{i_j} / A_n}{h_n} \right) \right. \right. \\
 & \quad \left. \left. \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right| \right\|_{\mathcal{F}_2 \mathcal{K}^2},
 \end{aligned}$$

and:

$$\begin{aligned}
 & \tilde{d}_{nh, 2}^{(1)} \left(\zeta_{1, K_{2,1}} W'(u), \zeta_{2, K_{2,2}} W''(u) \right) \\
 & := \mathbb{E}_\epsilon \left| \frac{1}{n^{3/2} h^d \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \epsilon_p \epsilon_q \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \left[\zeta_{1i_1} \zeta_{1i_2} \right. \right. \\
 & \quad \left. \left. \prod_{k=1}^2 K_{1,2} \left(\frac{d(x_k, \eta_{i_k})}{h} \right) W'_{s_i, A_n}(u) - \zeta_{2i_1} \zeta_{2i_2} \prod_{k=1}^2 K_{2,2} \left(\frac{d(x_k, \eta_{i_k})}{h} \right) W''_{s_i, A_n}(u) \right] \right|.
 \end{aligned}$$

Let consider another semi-norm $\tilde{d}_{nh, 2}^{(2)}$:

$$\begin{aligned}
 & \tilde{d}_{nh, 2}^{(2)} \left(\zeta_{1, K_{2,1}} W'(u), \zeta_{2, K_{2,2}} W''(u) \right) \\
 & = \frac{1}{nh^d \phi^2(h)} \left[\sum_{\ell_1 \neq \ell_2 \in L_n} \left(\zeta_{1i_1} \zeta_{1i_2} \prod_{k=1}^2 K_{1,2} \left(\frac{d(x_k, \eta_{i_k})}{h} \right) W'_{s_i, A_n}(u) \right. \right. \\
 & \quad \left. \left. - \zeta_{2i_1} \zeta_{2i_2} \prod_{k=1}^2 K_{2,2} \left(\frac{d(x_k, \eta_{i_k})}{h} \right) W''_{s_i, A_n}(u) \right)^2 \right]^{1/2}.
 \end{aligned}$$

One can see that

$$\tilde{d}_{nh, 2}^{(1)} \left(\zeta_{1, K_{2,1}} W'(u), \zeta_{2, K_{2,2}} W''(u) \right) \leq A_{1,n} n^{-1/2} h^d \phi(h) \tilde{d}_{nh, 2}^{(2)} \left(\zeta_{1, K_{2,1}} W'(u), \zeta_{2, K_{2,2}} W''(u) \right).$$

We readily infer that

$$E_{\cdot|S} \left\| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{i, \varphi, n}^{(u)} \right\|_{\mathcal{F}_2 \mathcal{H}^2} \tag{76}$$

$$\leq c_2 E_{\cdot|S} \int_0^{D_{nh}^{(U_1)}} N(t A_{1,n}^{-d} n^{1/2}, \mathcal{F}_{i,j}, \tilde{d}_{nh,2}^{(2)}) dt$$

$$\leq c_2 A_{1,n}^d n^{-1/2} \mathbb{P} \left\{ D_{nh}^{(U_1)} A_{1,n}^{-d} n^{1/2} \geq \lambda_n \right\} + c_m A_{1,n}^d n^{-1/2} \int_0^{\lambda_n} \log t^{-1} dt,$$

where $\lambda_n \rightarrow 0$. We have

$$\frac{\left(\int_0^{\lambda_n} \log t^{-1} dt \right)}{\left(\lambda_n \log \lambda_n^{-1} \right)} \rightarrow 0,$$

where λ_n must be chosen in such a way that the following relation will be achieved

$$A_{1,n}^d \lambda_n n^{-1/2} \log \lambda_n^{-1} \rightarrow 0. \tag{77}$$

By utilizing the triangle inequality in conjunction with Hoeffding’s trick, we are easily able to derive that

$$A_{1,n}^d n^{-1/2} \mathbb{P} \left\{ D_{nh}^{(U_1)} \geq \lambda_n A_{1,n}^d n^{-1/2} \right\}$$

$$\leq \lambda_n^{-2} A_{1,n}^{-d} n^{-5/2} h \phi^{-1}(h) E_{\cdot|S} \left\| \sum_{\ell_1 \neq \ell_2 \in L_n} \left[\sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \tilde{\zeta}_{i_1} \tilde{\zeta}_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_2 \left(\frac{d(x_2, \eta'_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right]^2 \right\|_{\mathcal{F}_2 \mathcal{H}^2} \tag{78}$$

$$\leq c_2 \llbracket L_n \rrbracket \lambda_n^{-2} A_{1,n}^{-d} n^{-5/2} h \phi^{-1}(h) E_{\cdot|S} \left\| \sum_{\ell_1 \in L_n} \left[\sum_{i_1, i_2: s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \tilde{\zeta}_{i_1} \tilde{\zeta}_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_2 \left(\frac{d(x_2, \eta'_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right]^2 \right\|_{\mathcal{F}_2 \mathcal{H}^2},$$

where $\{\eta'_i\}_{i \in \mathbb{N}^*}$ are independent copies of $(\eta_i)_{i \in \mathbb{N}^*}$. By imposing

$$\lambda_n^{-2} A_{1,n}^{d-r} n^{-1/2} \rightarrow 0, \tag{79}$$

we readily infer that

$$\left\| \llbracket L_n \rrbracket \lambda_n^{-2} A_{1,n}^{-d} n^{-5/2} h \phi^{-1}(h) E_{\cdot|S} \sum_{\ell_1 \in L_n} \left[\sum_{i_1, i_2: s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \tilde{\zeta}_{i_1} \tilde{\zeta}_{i_2} \prod_{k=1}^2 K_2 \left(\frac{d(x_k, \eta_{i_k})}{h} \right) W_{s_i, A_n}^{(u)} \right]^2 \right\|_{\mathcal{F}_2 \mathcal{H}^2}$$

$$\leq O \left(\lambda_n^{-2} A_{1,n}^{d-r} n^{-1/2} \right).$$

A symmetrization of the last inequality in (78) succeeded by an application of the Proposition A1 in the Appendix A gives

$$\begin{aligned}
 & \ll \left\| \left[\sum_{\ell_1 \in L_n} \left[\sum_{i_1, i_2: s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \epsilon_p \zeta_{i_1} \zeta_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) \right. \right. \right. \\
 & \qquad \left. \left. \left. K_2 \left(\frac{d(x_2, \eta'_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right] \right]^2 \right\|_{\mathcal{F}_2 \mathcal{H}^2} \tag{80} \\
 & \leq c_2 E_{\cdot | S} \left(\int_0^{D_{nh}^{(U_2)}} (\log N(u, \mathcal{F}_{i,j}, \tilde{d}_{nh,2}^j))^{1/2} \right),
 \end{aligned}$$

where

$$\begin{aligned}
 D_{nh}^{(U_2)} &= \left\| \mathbb{E}_\epsilon \left[\left[L_n \right] \lambda_n^{-2} A_{1,n}^{-d} n^{-5/2} \phi^{-1}(h) \right. \right. \\
 & \qquad \left. \left. \sum_{\ell_1 \in L_n} \left[\sum_{i_1, i_2: s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \zeta_{i_1} \zeta_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_2 \left(\frac{d(x_2, \eta'_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right] \right]^2 \right\|_{\mathcal{F}_2 \mathcal{H}^2}.
 \end{aligned}$$

and for $\zeta_{1, K_{2,1} W'}$, $\zeta_{2, K_{2,2} W''} \in \mathcal{F}_{ij}$:

$$\begin{aligned}
 & \tilde{d}_{nh,2}^j \left(\zeta_{1, K_{2,1} W'}^{(u)}, \zeta_{2, K_{2,2} W''}^{(u)} \right) \\
 & := \mathbb{E}_\epsilon \left[\left[L_n \right] \lambda_n^{-2} A_{1,n}^{-d} n^{-5/2} \phi^{-1}(h) \sum_{\ell_1 \in L_n} \epsilon_p \left[\left(\sum_{i_1, i_2: s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \zeta_{1i_1} \zeta_{1i_2} K_{2,1} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) \right. \right. \right. \\
 & \qquad \left. \left. \left. K_{2,1} \left(\frac{d(x_2, \eta'_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right)^2 - \left(\sum_{i_1, i_2 \in H_p^{(u)}} \zeta_{2i} \zeta_{2j} K_{2,2} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_{2,2} \left(\frac{d(x_2, \eta'_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right)^2 \right] \right].
 \end{aligned}$$

By the fact that:

$$\begin{aligned}
 & \mathbb{E}_\epsilon \left[\left[L_n \right] \lambda_n^{-2} A_{1,n}^{-d} n^{-5/2} \phi^{-1}(h) \sum_{\ell_1 \in L_n} \epsilon_p \left(\sum_{i_1, i_2: s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \zeta_{i_1} \zeta_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) \right. \right. \\
 & \qquad \left. \left. K_2 \left(\frac{d(x_2, \eta'_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right)^2 \right] \\
 & \leq A_{1,n}^{3d/2} \lambda_n^{-2} n^{-1} \left[\left[L_n \right]^{-1} A_{1,n}^{-2d} \phi^{-2}(h_n) \sum_{\ell_1 \in L_n} \sum_{i_1, i_2: s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \left(\zeta_{i_1} \zeta_{i_2} K_2 \left(\frac{d(x_i, \eta_{i_1})}{h} \right) \right. \right. \\
 & \qquad \left. \left. K_2 \left(\frac{d(x_2, \eta'_j)}{h} \right) W_{s_i, A_n}^{(u)} \right)^4 \right]^{1/2},
 \end{aligned}$$

so:

$$A_{1,n}^{d3/2} \lambda_n^{-2} n^{-1} \rightarrow 0, \tag{81}$$

we have the convergence of (80) to zero. Recall that

$$\left[L_n \right] = O \left((A_n / A_{3,n})^d \right) \lesssim (A_n / A_{1,n})^d.$$

(II): The Same Block

$$\begin{aligned}
 & \mathbb{P} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \right. \right. \\
 & \quad \left. \left. \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) \right\} W_{s_i, A_n} \right| > \delta \right) \\
 & \leq \mathbb{P} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \right. \right. \\
 & \quad \left. \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \left[\prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) \right] W_{s_i, A_n} \right| > \delta \right) \\
 & \quad + \mathbb{P} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \right. \right. \\
 & \quad \left. \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) \left[W_{s_i, A_n} - W_{s_i, A_n}^{(u)} \right] \right| > \delta \right) \\
 & \quad + \mathbb{P} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \right. \right. \\
 & \quad \left. \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, X_{u_j}(s_{i_j}))}{h} \right) W_{s_i, A_n}^{(u)} \right| > \delta \right) \\
 & \leq \mathbb{P} \left(\sup_{\mathcal{F}_m \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \right. \right. \\
 & \quad \left. \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right| > \delta \right) \\
 & \quad + C \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d) + o_{\mathbb{P}}(1) + o_{\mathbb{P}}(1),
 \end{aligned}$$

In the same manner as I, we can show that the first and the second term in the previous inequality is of order $o_{\mathbb{P}}(1)$. So, as the preceding proof, it suffices to prove that

$$E_{\cdot|S} \left(\left\| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \right\| \right)$$

$$\prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \Big\|_{\mathcal{F}_2, \mathcal{H}^2} \rightarrow 0.$$

Notice that we treat uniformly bounded classes functions in which we obtain uniformly in $B^m \times \mathcal{F}_2 \mathcal{H}^2$

$$E_{\cdot|S} \left(\sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right) = O(a_n).$$

This implies that we have to prove that, for $u \in B^m$

$$\begin{aligned} & \mathbb{E} E_{\cdot|S} \left(\left\| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_1 \neq i_2 \\ i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n}} \right. \right. \\ & \left. \left[\prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right. \right. \\ & \left. \left. - E_{\cdot|S} \left(\prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right) \right] \right\|_{\mathcal{F}_2, \mathcal{H}^2} \right) \rightarrow 0. \end{aligned} \tag{82}$$

As for empirical processes, to prove (82), it is enough to symmetrize and show that

$$\begin{aligned} & E_{\cdot|S} \left\| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \epsilon_p \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \right. \\ & \left. \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2} \rightarrow 0. \end{aligned}$$

Similarly to how in (75), we have

$$\begin{aligned} & E_{\cdot|S} \left\| \frac{1}{n^{3/2} h^{d+1} \phi(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \epsilon_p \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \right. \\ & \left. \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2} \leq \mathbb{E} \left(\int_0^{D_{nh}^{(U_3)}} (\log N(u, \mathcal{F}_{i_1, i_2}, \tilde{d}_{nh, 2}^{(3)}))^{1/2} du \right), \end{aligned}$$

where

$$D_{nh}^{(U_3)} = \left\| \mathbb{E}_\epsilon \left\| \frac{1}{n^{3/2} h^d \phi(h)} \sum_{\ell_1 \in L_n} \epsilon_p \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \right. \right.$$

$$\prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \Bigg\|_{\mathcal{F}_2, \mathcal{H}^2}, \tag{83}$$

and the semi-metric $\tilde{d}_{nh,2}^{(3)}$ is defined by

$$\begin{aligned} & \tilde{d}_{nh,2}^{(3)} \left(\xi_{1, K_{2,1}} W^{(u)}, \xi_{2, K_{2,2}} W''^{(u)} \right) \\ &= \mathbb{E}_\epsilon \left| \frac{1}{n^{3/2} h^d \phi(h)} \sum_{\ell_1 \in L_n} \epsilon_p \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \left(\xi_{1i_1} \xi_{1j} K_{2,1} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) \right. \right. \\ & \quad \left. \left. K_{2,1} \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} - \xi_{2i_2} \xi_{2j} K_{2,2} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_{2,2} \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W''^{(u)} \right) \right|. \end{aligned}$$

Since we are considering uniformly bounded classes of functions, we obtain

$$\begin{aligned} & \mathbb{E}_\epsilon \left| n^{-3/2} h \phi^{-1}(h_n) \sum_{\ell_1 \in L_n} \epsilon_p \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \xi_{i_1} \xi_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) \right. \\ & \quad \left. K_2 \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right| \\ & \leq A_{1,n}^{3d/2} n^{-1} h \phi^{-1}(h_n) \left[\frac{1}{\|L_n\| A_{1,n}^2} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \left(\xi_{i_1} \xi_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) \right. \right. \\ & \quad \left. \left. K_2 \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W_{s_i, A_n}^{(u)} \right)^2 \right]^{1/2} \leq O \left(A_{1,n}^{3d/2} n^{-1} \phi^{-1}(h_n) \right). \end{aligned}$$

Since $A_{1,n}^{3d/2} n^{-1} \phi^{-1}(h) \rightarrow 0, D_{nh}^{(U_3)} \rightarrow 0$, we obtain $\Pi \rightarrow 0$ as $n \rightarrow \infty$.

(III): Different Types of Blocks

Avoiding the repetition, we can directly see that:

$$\begin{aligned} & \mathbb{P} \left(\sup_{\mathcal{F}_m, \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \right. \right. \\ & \quad \left. \left. \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) \right\} W_{s_i, A_n} \right| > \delta \right) \\ & \leq \mathbb{P} \left(\sup_{\mathcal{F}_m, \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \frac{1}{n^{3/2} h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \right. \right. \\ & \quad \left. \left. \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{u_j - s_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{s_i, A_n}^{(u)} \right| > \delta \right) \tag{84} \end{aligned}$$

$$+C\left(\frac{A_n}{A_{1,n}}\right)^d \beta\left(A_{2,n}; A_n^d\right) + o_{\mathbb{P}}(1) + o_{\mathbb{P}}(1).$$

For $p = 1$ and $p = v_n$:

$$\begin{aligned} & E_{\cdot|S} \left\| \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \right. \\ & \quad \left. \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^2 K_2\left(\frac{d(x_j, \eta_{i_j})}{h}\right) W_{\mathbf{s}_{i,A_n}}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2} \\ &= E_{\cdot|S} \left\| \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i}=3}} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \right. \\ & \quad \left. \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^2 K_2\left(\frac{d(x_j, \eta_{i_j})}{h}\right) W_{\mathbf{s}_{i,A_n}}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2}. \end{aligned}$$

For $2 \leq p \leq v_n - 1$, we obtain:

$$\begin{aligned} & E_{\cdot|S} \left\| \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \right. \\ & \quad \left. \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^2 K_2\left(\frac{d(x_j, \eta_{i_j})}{h}\right) W_{\mathbf{s}_{i,A_n}}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2} \\ &= E_{\cdot|S} \left\| \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i}=4}} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \right. \\ & \quad \left. \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^2 K_2\left(\frac{d(x_j, \eta_{i_j})}{h}\right) W_{\mathbf{s}_{i,A_n}}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2} \\ &\leq E_{\cdot|S} \left\| \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i}=3}} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \right. \\ & \quad \left. \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^2 K_2\left(\frac{d(x_j, \eta_{i_j})}{h}\right) W_{\mathbf{s}_{i,A_n}}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2}, \end{aligned}$$

therefore, it suffices to show that:

$$\begin{aligned} & E_{\cdot|S} \left\| \frac{1}{n^{3/2}h^{2d}\phi^2(h)} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i}=3}} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \right. \\ & \quad \left. \prod_{j=1}^2 \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \prod_{j=1}^2 K_2\left(\frac{d(x_j, \eta_{i_j})}{h}\right) W_{\mathbf{s}_{i,A_n}}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2}. \end{aligned} \tag{85}$$

By similar arguments as in [82], the usual symmetrization is applied and:

$$\begin{aligned}
 & E_{\cdot|S} \left\| \frac{\llbracket L_n \rrbracket}{n^{3/2} h^{2d} \phi^2(h)} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3 \\ \ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \right. \\
 & \quad \left. \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_{i_j}, A_n}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{K}^2} \\
 & \leq 2E_{\cdot|S} \left\| \frac{\llbracket L_n \rrbracket}{n^{3/2} h^{2d} \phi^2(h)} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3 \\ \ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \epsilon_q \right. \\
 & \quad \left. \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_{i_j}, A_n}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{K}^2} \\
 & = 2E_{\cdot|S} \left\{ \left\| \frac{\llbracket L_n \rrbracket}{n^{3/2} h^{2d} \phi^2(h)} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3 \\ \ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \epsilon_q \right. \right. \\
 & \quad \left. \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_{i_j}, A_n}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{K}^2} \mathbf{1}_{\{D_{nh}^{(U_4)} \leq \gamma_n\}} \right\} \tag{86} \\
 & + 2E_{\cdot|S} \left\{ \left\| \frac{\llbracket L_n \rrbracket}{n^{3/2} h^{2d} \phi^2(h)} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3 \\ \ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \right. \right. \\
 & \quad \left. \left. \epsilon_q \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_{i_j}, A_n}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{K}^2} \mathbf{1}_{\{D_{nh}^{(U_4)} > \gamma_n\}} \right\} \\
 & = 2\text{III}_1 + 2\text{III}_2,
 \end{aligned}$$

where

$$\begin{aligned}
 & D_{nh}^{(U_4)} \\
 & = \left\| \frac{\llbracket L_n \rrbracket}{n^{3/2} h^{2d} \phi^2(h)} \left[\sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3 \\ \epsilon \neq \epsilon_0}} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \left(\sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \right. \right. \right. \\
 & \quad \left. \left. \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_{i_j}, A_n}^{(u)} \right)^2 \right]^{1/2} \right\|_{\mathcal{F}_2, \mathcal{K}^2}. \tag{87}
 \end{aligned}$$

In a similar way as in (75), we infer that

$$\text{III}_1 \leq c_2 \int_0^{\gamma_n} \left(\log N \left(t, \mathcal{F}_{i_1, i_2}, \tilde{d}_{nh,2}^{(4)} \right) \right)^{1/2} dt, \tag{88}$$

where

$$\tilde{d}_{nh,2}^{(4)} \left(\tilde{\zeta}_1, K_{2,1} W^{(u)}, \tilde{\zeta}_2, K_{2,2} W''^{(u)} \right)$$

$$:= \mathbb{E}_\epsilon \left| \llbracket L_n \rrbracket n^{-3/2} h \phi^{-1}(h_n) \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \right. \\ \left. \epsilon_q \left[\zeta_{1i_1} \zeta_{1i_2} K_{2,1} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_{2,1} \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W'_{s_i, A_n}(u) \right. \right. \\ \left. \left. - \zeta_{2i_1} \zeta_{2i_2} K_{2,2} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_{2,2} \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W''_{s_i, A_n}(u) \right] \right|.$$

Since we have

$$\mathbb{E}_\epsilon \left| \llbracket L_n \rrbracket n^{3/2} h^{2d} \phi^2(h) \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \epsilon_q \zeta_{i_1} \zeta_{i_2} \right. \\ \left. K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_2 \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W_{s_i, A_n}(u) \right| \\ \leq A_{1,n}^{-d/2} A_{2,n}^d h^{-d+1} \phi(h) \left(\frac{1}{A_{1,n}^d A_{2,n}^d \llbracket L_n \rrbracket h^{d-1} \phi^4(h_n)} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3} \right. \\ \left. \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \left[\zeta_{i_1} \zeta_{i_2} K_2 \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_2 \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W_{s_i, A_n}(u) \right]^2 \right)^{1/2},$$

and considering the semi-metric

$$\tilde{d}_{nh,2}^{(5)} \left(\zeta_{1 \cdot} K_{2,1} W'(u), \zeta_{2 \cdot} K_{2,2} W''(u) \right) \\ := \left(\frac{1}{A_{1,n}^d A_{2,n}^d \llbracket L_n \rrbracket h^{d-1} \phi^4(h_n)} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \left[\zeta_{1i_1} \zeta_{1i_2} \right. \right. \\ \left. \left. K_{2,1} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_{2,1} \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W'_{s_i, A_n}(u) - \zeta_{2i_1} \zeta_{2i_2} K_{2,2} \left(\frac{d(x_1, \eta_{i_1})}{h} \right) K_{2,2} \left(\frac{d(x_2, \eta_{i_2})}{h} \right) W''_{s_i, A_n}(u) \right]^2 \right)^{1/2}.$$

We demonstrate that the statement in (88) is bounded as follows

$$\llbracket L_n \rrbracket^{1/2} A_{2,n}^d n^{-1/2} h^2 \phi(h) \int_0^{\llbracket L_n \rrbracket^{-1/2} A_{2,n}^{-d} n^{1/2} h^{2d} \gamma_n} \left(\log N \left(t, \mathcal{F}_{i_1, i_2}, \tilde{d}_{nh,2}^{(5)} \right) \right)^{1/2} dt,$$

by choosing $\gamma_n = n^{-\alpha}$ for some $\alpha > (17r - 26)/60r$, we obtain the convergence of the preceding quantity to zero. In order to bound the second term on the right-hand side of (86), we can mention that

$$\text{III}_2 = \mathbb{E} \left\{ \left| \llbracket L_n \rrbracket n^{-3/2} h \phi^{-2}(h) \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \epsilon_q \right. \right.$$

$$\begin{aligned}
 & \left. \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_i, A_n}^{(u)} \right\|_{\mathcal{F}_2, \mathcal{H}^2} \mathbf{1}_{\{D_{nh}^{(U_4)} > \gamma_n\}} \quad (89) \\
 \leq & A_{1,n}^{-1} A_{2,n} n^{1/2} h^d \phi^{-1}(h) \mathbb{P} \left\{ \left\| \left[L_n \right]^2 n^{-3} h^2 \phi^{-2}(h_n) \sum_{\substack{\ell_2: \min_{1 \leq i \leq d} \ell_{2i} = 3 \\ \ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \left(\sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_i, A_n}^{(u)} \right)^2 \right\|_{\mathcal{F}_2, \mathcal{H}^2} \geq \gamma_n^2 \right\}.
 \end{aligned}$$

We are going to use the square root method on the last expression conditionally on $\Gamma_n(\ell_1; \epsilon_0) \cap R_n$. We denote by $E_{\epsilon \neq \epsilon_0}$ the expectation with respect to $\sigma\{\eta_{i_2}, \epsilon \neq \epsilon_0\}$ and we will suppose that any class of functions \mathcal{F}_m is unbounded and its envelope function satisfies for some $p > 2$:

$$\theta_p := \sup_{x \in \mathcal{F}_m} \mathbb{E}(F^p(\mathbf{Y}) | \mathbf{X} = x) < \infty, \tag{90}$$

for $2r/(r - 1) < s < \infty$ (in the notation in of [151] [Lemma 5.2]).

$$\begin{aligned}
 M_n = & \left[L_n \right]^{1/2} E_{\epsilon \neq \epsilon_0} \left(\sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \prod_{j=1}^2 \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) \right. \\
 & \left. \prod_{j=1}^2 K_2 \left(\frac{d(x_j, \eta_{i_j})}{h} \right) W_{\mathbf{s}_i, A_n}^{(u)} \right)^2,
 \end{aligned}$$

where

$$x = \gamma_n^2 A_{1,n}^{5d/2} n^{1/2} h^{md/2} \phi^{-m/2}(h), \quad \rho = \lambda = 2^{-4} \gamma_n A_{1,n}^{5d/4} n^{1/4} h^{md/4} \phi^{-m/4}(h),$$

and

$$m = \exp \left(\gamma_n^2 n h^{2d} \phi^{-2}(h_n) A_{2,n}^{-2d} \right).$$

However, since we need $t > 8M_n$, and $m \rightarrow \infty$, by similar arguments as in ([82], p. 69), we reach the convergence of (88) and (89) to zero.

(IV): Blocks of Different Types

The target here is to prove that:

$$\begin{aligned}
 \mathbb{P} \left(\sup_{\mathcal{F}_m, \mathcal{H}^m} \sup_{x \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \left| \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_1} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \right. \right. \\
 \left. \left. \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{\mathbf{s}_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_i, A_n} \right| > \delta \right) \rightarrow 0.
 \end{aligned}$$

We have

$$\begin{aligned} & \left\| n^{-3/2} \frac{1}{h^{2d} \phi^2(h)} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_1} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \right. \\ & \quad \left. \times \prod_{j=1}^2 \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\} W_{\mathbf{s}_{i_j}, A_n} \right\|_{\mathcal{F}_2 \mathcal{K}^2} \\ & \leq c_2 \llbracket L_n \rrbracket A_{1,n}^d A_{2,n}^d n^{-3/2} h^{-d} \phi^{-1}(h) \rightarrow 0. \end{aligned}$$

Hence, the proof of the lemma is complete.

The final step in the proof of Proposition 1 lies in the use of Lemma 1 to prove that the nonlinear term converges to zero. \square

9.2.2. Proof of Theorem 1

We have

$$\begin{aligned} & \widehat{r}_n^{(m)}(\varphi, \mathbf{x}, \mathbf{u}; h_n) - r^{(m)}(\varphi, \mathbf{x}, \mathbf{u}) \\ & = \frac{1}{\widetilde{r}_1(\varphi, \mathbf{x}, \mathbf{u})} \left(\widehat{g}_1(\mathbf{u}, \mathbf{x}) + \widehat{g}_2(\mathbf{u}, \mathbf{x}) - r^{(m)}(\varphi, \mathbf{x}, \mathbf{u}) \widetilde{r}_1(\varphi, \mathbf{x}, \mathbf{u}) \right), \end{aligned}$$

where

$$\begin{aligned} \widetilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}) & = \frac{(n-m)!}{n! h^{md} \phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}, \\ \widehat{g}_1(\mathbf{u}, \mathbf{x}) & = \frac{(n-m)!}{n! h^{md} \phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\} \prod_{j=1}^m \epsilon_{s_{i_j}, A_n}, \\ \widehat{g}_2(\mathbf{u}, \mathbf{x}) & = \frac{(n-m)!}{n! h^{md} \phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \bar{K} \left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j} / A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\} \\ & \quad \times r^{(m)} \left(\varphi, \mathbf{X}_{s_{i_1}, A_n}, \dots, \mathbf{X}_{s_{i_m}, A_n}, \frac{\mathbf{s}_{i_1}}{A_n}, \dots, \frac{\mathbf{s}_{i_m}}{A_n} \right). \end{aligned}$$

The proof of this theorem is involved and divided into the following four steps, where in each one, we aim to show that

Step 1.

$$\sup_{\mathcal{F}_m \mathcal{K}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} |\widehat{g}_1(\mathbf{u}, \mathbf{x})| = O_{\mathbb{P}} \left(\sqrt{\log n / n h^{md} \phi^m(h)} \right).$$

Step 2.

$$\begin{aligned} & \sup_{\mathcal{F}_m \mathcal{K}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} |\widehat{g}_2(\mathbf{u}, \mathbf{x}) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \widetilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}; h_n)| \\ & - E_{\cdot|S} (\widehat{g}_2(\mathbf{u}, \mathbf{x}) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \widetilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}; h_n)) = O_{\mathbb{P}} \left(\sqrt{\log n / n h^{md} \phi^m(h)} \right). \end{aligned}$$

Step 3. Let $\kappa_2 = \int_{\mathbb{R}} x^2 K(x) dx$.

$$\sup_{\mathcal{F}_m \mathcal{K}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in B^m} \left| E_{\cdot|S} \left(\widehat{g}_2(\mathbf{u}, \mathbf{x}) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \widetilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}; h_n) \right) \right|$$

$$= O\left(\frac{1}{A_n^{dp} \phi(h)}\right) + o(h^2), P_S - a.s.$$

Step 4.

$$\sup_{\mathcal{F}_m} \sup_{\mathcal{K}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \tilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}) - E_{\cdot|S}(\tilde{r}_1(\varphi, \mathbf{u}, \mathbf{x})) \right| = o_{P_{\cdot|S}}(1).$$

It is clear that **Step 1** follows directly from Proposition 1 for $W_{s_i, A_n} = \prod_{j=1}^m \epsilon_{s_{i_j}, A_n}$. The second one (**Step 2**) holds also if we replace W_{s_i, A_n} with $\hat{g}_2(\mathbf{u}, \mathbf{x}) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \tilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}; h_n)$ then applying Proposition 1.

We will pass now to **Step 4**. Observe that for $W_{s_i, A_n} \equiv 1$, the previous mentioned proposition proved that

$$\sup_{\mathcal{F}_m} \sup_{\mathcal{K}^m} \sup_{x \in \mathcal{H}^m} \sup_{u \in B^m} \left| \tilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}) - E_{\cdot|S}(\tilde{r}_1(\varphi, \mathbf{u}, \mathbf{x})) \right| = o_{P_{\cdot|S}}(1).$$

Step 3. will be treated in what follows:

Let $K_0 : [0, 1] \rightarrow \mathbb{R}$ be a Lipschitz continuous function compactly support on $[-qC_1, qC_1]$ for some $q > 1$ and such that $K_0(x) = 1, \forall x \in [-C_1, C_1]$. Show that

$$E_{\cdot|S} \left[\hat{g}_2(\mathbf{u}, \mathbf{x}) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \tilde{r}_1(\varphi, \mathbf{u}, \mathbf{x}; h_n) \right] = \sum_{i=1}^4 Q_i(\mathbf{u}, \mathbf{x}), \tag{91}$$

where Q_i can be defined as follows

$$Q_i(\mathbf{u}, \mathbf{x}) = \frac{(n-m)!}{n! h^{md} \phi^m(h)} \sum_{i \in I_n^m} \left\{ \prod_{j=1}^m \bar{K} \left(\frac{\mathbf{u}_j - s_{i_j} / A_n}{h_n} \right) \right\} q_i(\mathbf{u}, \mathbf{x}), \tag{92}$$

such that

$$\begin{aligned} q_1(\mathbf{u}, \mathbf{x}) &= E_{\cdot|S} \left[\prod_{j=1}^m K_0 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) \left\{ \prod_{j=1}^m K_2 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) \right. \right. \\ &\quad \left. \left. - \prod_{j=1}^m K_2 \left(\frac{d \left(x_j, X_{\frac{s_{i_j}}{A_n}}(s_{i_j}) \right)}{h} \right) \right\} \times \left\{ r^{(m)} \left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{s_i, A_n} \right) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \right\} \right], \\ q_2(\mathbf{u}, \mathbf{x}) &= E_{\cdot|S} \left[\prod_{j=1}^m \left\{ K_0 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) K_2 \left(\frac{d \left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j}) \right)}{h} \right) \right\} \right. \\ &\quad \left. \left\{ r^{(m)} \left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{s_i, A_n} \right) - r^{(m)} \left(\varphi, \frac{\mathbf{s}_i}{A_n}, X_{s_i / A_n}(s_i) \right) \right\} \right], \\ q_3(\mathbf{u}, \mathbf{x}) &= E_{\cdot|S} \left[\left\{ \prod_{j=1}^m K_0 \left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n} \right) - \prod_{j=1}^m K_0 \left(\frac{d \left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j}) \right)}{h} \right) \right\} \right] \end{aligned}$$

$$q_4(\mathbf{u}, \mathbf{x}) = E_{\cdot|S} \left[\prod_{j=1}^m K_2 \left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h} \right) \times \left\{ r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{\mathbf{s}_i/A_n}(s_i)\right) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \right\} \right],$$

$$= E_{\cdot|S} \left[\prod_{j=1}^m K_2 \left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h} \right) \left\{ r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{\mathbf{s}_i/A_n}(s_i)\right) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \right\} \right].$$

Observe that

$$Q_1(\mathbf{u}, \mathbf{x}) \lesssim \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \left\{ \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) E_{\cdot|S} \left[\prod_{j=1}^m K_0\left(\frac{d(x_j, X_{\mathbf{s}_{i_j}/A_n})}{h_n}\right) \right. \right.$$

$$\left. \left| \prod_{j=1}^m K_2\left(\frac{d(x_j, X_{\mathbf{s}_{i_j}/A_n})}{h_n}\right) - \prod_{j=1}^m K_2\left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h}\right) \right| \right]$$

$$\left. \times \left| r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{\mathbf{s}_i/A_n}\right) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \right| \right\},$$

using the properties of $r^{(m)}(\mathbf{u}, \mathbf{x})$ allows us to show that

$$\prod_{j=1}^m K_0\left(\frac{d(x_j, X_{\mathbf{s}_{i_j}/A_n})}{h_n}\right) \left| r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{\mathbf{s}_i/A_n}\right) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \right| \leq Ch^m$$

$$Q_1(\mathbf{u}, \mathbf{x}) \leq \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \left\{ \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \right\} E_{\cdot|S} [Ch^m$$

$$\times C \sum_{j=1}^m \left| K_2\left(\frac{d(x_j, X_{\mathbf{s}_{i_j}/A_n})}{h_n}\right) - K_2\left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h}\right) \right|^p] \quad (\text{Using the telescoping$$

argument, and the boundness of K_2 for $p = \min(\rho, 1)$ and $C < \infty$)

$$\leq \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \left\{ \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{s}_{i_j}/A_n}{h_n}\right) \right\} E_{\cdot|S} \left[Ch^m \sum_{j=1}^m \left| \frac{C}{A_n^d h} U_{\mathbf{s}_{i_j}/A_n}\left(\frac{\mathbf{s}_{i_j}}{A_n}\right) \right|^p \right]$$

$$\leq \frac{C}{A_n^{pd}\phi^m(h)h^{p-m}} \quad \text{uniformly in } \mathbf{u}.$$

In a similar way, and for

$$\mathbb{E} \left[\prod_{j=1}^m K_2\left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h}\right) \right] \leq C\phi^{m-1}(h),$$

and since $r^{(m)}(\cdot)$ is Lipschitz and $d\left(\mathbf{X}_{s_{i_j}, A_n}, \mathbf{X}_{\frac{s_{i_j}}{A_n}}(s_j)\right) \leq \frac{C}{A_n^d} U_{s_{i_j}, A_n}\left(\frac{s_{i_j}}{A_n}\right)$ and the variable $U_{s_{i_j}, A_n}\left(\frac{s_{i_j}}{A_n}\right)$ has a finite p -th moment, we can see that

$$\begin{aligned}
 Q_2(\mathbf{u}, \mathbf{x}) &= \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - s_{i_j}/A_n}{h_n}\right) \\
 &\quad E_{\cdot|S} \left[\prod_{j=1}^m \left\{ K_0\left(\frac{d(x_j, X_{s_{i_j}, A_n})}{h_n}\right) K_2\left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h}\right) \right\} \right. \\
 &\quad \left. \left\{ r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{s_{i_j}, A_n}\right) - r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{s_i/A_n}(s_i)\right) \right\} \right] \\
 &\leq \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \left\{ \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - s_{i_j}/A_n}{h_n}\right) \right\} E_{\cdot|S} \left[\phi^{m-1}(h) \left| \frac{C}{A_n^d} U_{s_{i_j}, A_n}\left(\frac{s_{i_j}}{A_n}\right) \right|^p \right] \\
 &\leq \frac{C}{A_n^{pd}\phi(h)},
 \end{aligned} \tag{93}$$

and

$$\sup_{\mathcal{F}_m \cdot \mathcal{H}^m} \sup_{\mathbf{x} \in \mathcal{H}^m} \sup_{\mathbf{u} \in I_h^m} Q_3(\mathbf{u}, \mathbf{x}) \lesssim \frac{1}{A_n^{pd}\phi^m(h)h^{p-m}}. \tag{94}$$

For the last term, we have

$$\begin{aligned}
 Q_4(\mathbf{u}, \mathbf{x}) &= \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - s_{i_j}/A_n}{h_n}\right) \\
 &\quad E_{\cdot|S} \left[\prod_{j=1}^m K_2\left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h}\right) \left\{ r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{s_i/A_n}(s_i)\right) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \right\} \right].
 \end{aligned}$$

Using Lemma A1 and inequality (17) and under Assumption 1, it follows that

$$\begin{aligned}
 |Q_4(\mathbf{u}, \mathbf{x})| &\leq \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - s_{i_j}/A_n}{h_n}\right) \\
 &\quad E_{\cdot|S} \left[\prod_{j=1}^m K_2\left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h}\right) \left| r^{(m)}\left(\varphi, \frac{\mathbf{s}_i}{A_n}, \mathbf{X}_{s_i/A_n}(s_i)\right) - r^{(m)}(\varphi, \mathbf{u}, \mathbf{x}) \right| \right] \\
 &\lesssim \frac{(n-m)!}{n!h^{md}\phi^m(h)} \sum_{i \in I_n^m} \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - s_{i_j}/A_n}{h_n}\right) \\
 &\quad E_{\cdot|S} \left[\prod_{j=1}^m K_2\left(\frac{d\left(x_i, X_{\frac{s_{i_j}}{A_n}}(s_{i_j})\right)}{h}\right) \left| d_{\mathcal{H}^m}(\mathbf{X}_{s_i/A_n}(s_i), \mathbf{x}) + \left\| \mathbf{u} - \frac{\mathbf{s}_i}{A_n} \right\|^\alpha \right| \right]
 \end{aligned} \tag{95}$$

$$\begin{aligned}
 &\lesssim \frac{(n-m)!}{n!h^{md}}\phi^m(h) \sum_{i \in I_n^m} \left| \prod_{j=1}^m \bar{K} \left(\frac{u_j - s_{i_j}/A_n}{h_n} \right) \right. \\
 &\quad \left. - \int_0^1 \cdots \int_0^1 \frac{1}{h^m} \prod_{j=1}^m \bar{K} \left(\frac{(u_j - v_j)}{h} \right) dv_j \right|_{E_{\cdot|S}} \left[\left| \prod_{j=1}^m K_2 \left(\frac{d \left(x_{i_j}, X_{s_{i_j}}(s_{i_j}) \right)}{h} \right) \right| \times h^\alpha \right] \\
 &\quad + \frac{(n-m)!}{n!h^{md}}\phi^m(h) \sum_{i \in I_n^m} \int_0^1 \cdots \int_0^1 \frac{1}{h^{md}} \prod_{j=1}^m \bar{K} \left(\frac{u_j - v_j}{h} \right) dv_j \times E_{\cdot|S} \left[\phi^{m-1}(h) h^\alpha \right] \\
 &\lesssim O_{P_{\cdot|S}} \left(h^{2 \wedge \alpha} \right).
 \end{aligned}$$

Adding the obtained results of $Q_i, 1 \leq i \leq 4$, **Step 3** yields the rate of convergence of the estimator. \square

9.2.3. Proof of Theorem 2

Recall that

$$\hat{r}_n^{(m)}(\varphi, \mathbf{x}, \mathbf{u}; h_n) = \frac{\sum_{i \in I_n^m} \varphi(\mathbf{Y}_{s_{i_1}, A_n}, \dots, \mathbf{Y}_{s_{i_m}, A_n}) \prod_{j=1}^m \left\{ \bar{K} \left(\frac{u_j - s_{i_j}/A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}}{\sum_{i \in I_n^m} \prod_{j=1}^m \left\{ \bar{K} \left(\frac{u_j - s_{i_j}/A_n}{h_n} \right) K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}}.$$

For $\mathbf{x} \in \mathcal{H}^m, \mathbf{y} \in \mathcal{Y}^m$, define

$$G_{\varphi, i}(\mathbf{x}, \mathbf{y}) := \frac{\prod_{j=1}^m \left\{ K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\} \varphi(\mathbf{Y}_{s_{i_1}, A_n}, \dots, \mathbf{Y}_{s_{i_m}, A_n})}{\mathbb{E} \prod_{j=1}^m \left\{ K_2 \left(\frac{d(x_j, \mathbf{X}_{s_{i_j}, A_n})}{h_n} \right) \right\}};$$

$$\mathcal{G} := \{ G_{\varphi, i}(\cdot, \cdot) \mid \varphi \in \mathcal{F}_m, \mathbf{i} = (i_1, \dots, i_m) \};$$

$$\mathcal{G}^{(k)} := \{ \pi_{k,m} G_{\varphi, i}(\cdot, \cdot), \varphi \in \mathcal{F}_m \};$$

$$\mathbb{U}_n(\varphi) = \mathbb{U}_n^{(m)}(G_{\varphi, i}) := \frac{(n-m)!}{n!} \sum_{i \in I_n^m} \prod_{j=1}^m \zeta_{i_j} G_{\varphi, i}(\mathbf{X}_{i_j}, \mathbf{Y}_{i_j});$$

and the U -empirical process is defined to be

$$\mu_n(\varphi) := \sqrt{nh^m \phi(h)} \{ \mathbb{U}_n(\varphi) - \mathbb{E}(\mathbb{U}_n(\varphi)) \}.$$

Then

$$\hat{r}_n^{(m)}(\varphi, \mathbf{x}, \mathbf{u}; h_n) = \frac{\mathbb{U}_n(\varphi)}{\mathbb{U}_n(1)}.$$

In order to establish the weak convergence of our estimator, it must be established first for $\mu_n(\varphi)$. We have mentioned before that we deal with unbounded classes of functions; that is why we should truncate the function $G_{\varphi, i}(\mathbf{x}, \mathbf{y})$, indeed, for $\lambda_n = n^{1/p}$, with $p > 0$, we have:

$$G_{\varphi, i}(\mathbf{x}, \mathbf{y}) = G_{\varphi, i}(\mathbf{x}, \mathbf{y}) \mathbf{1}_{\{F(\mathbf{y}) \leq \lambda_n\}} + G_{\varphi, i}(\mathbf{x}, \mathbf{y}) \mathbf{1}_{\{F(\mathbf{y}) > \lambda_n\}}$$

$$:= G_{\varphi,i}^{(T)}(\mathbf{x}, \mathbf{y}) + G_{\varphi,i}^{(R)}(\mathbf{x}, \mathbf{y}).$$

We can write the U -statistic as follows:

$$\begin{aligned} \mu_n(\varphi) &= \sqrt{nh^m\phi(h)} \left\{ \mathbb{U}_n^{(m)}\left(G_{\varphi,i}^{(T)}\right) - \mathbb{E}\left(\mathbb{U}_n^{(m)}\left(G_{\varphi,i}^{(T)}\right)\right) \right\} \\ &\quad + \sqrt{nh^m\phi(h)} \left\{ \mathbb{U}_n^{(m)}\left(G_{\varphi,i}^{(R)}\right) - \mathbb{E}\left(\mathbb{U}_n^{(m)}\left(G_{\varphi,i}^{(R)}\right)\right) \right\} \\ &:= \sqrt{nh^m\phi(h)} \left\{ \mathbb{U}_n^{(T)}(\varphi, \mathbf{i}) - \mathbb{E}\left(\mathbb{U}_n^{(T)}(\varphi)\right) \right\} \\ &\quad + \sqrt{nh^m\phi(h)} \left\{ \mathbb{U}_n^{(R)}(\varphi) - \mathbb{E}\left(\mathbb{U}_n^{(R)}(\varphi)\right) \right\} \\ &:= \mu_n^{(T)}(\varphi) + \mu_n^{(R)}(\varphi). \end{aligned} \tag{96}$$

The first term is the truncated part and the second is the remaining one. We have to prove that:

1. $\mu_n^{(T)}(\varphi)$ converges to a Gaussian process.
2. The remainder part does not matter much, in the sense that

$$\left\| \sqrt{nh^m\phi(h)} \left\{ \mathbb{U}_n^{(R)}(\varphi) - \mathbb{E}\left(\mathbb{U}_n^{(R)}(\varphi)\right) \right\} \right\|_{\mathcal{F}_m \mathcal{X}^m} \xrightarrow{\mathbb{P}} 0.$$

For the first point, we will use the decomposition of Hoeffding, which would be the same as the previous decomposition in Section 3.1 except that we replace $W_{i,n}$ by $\varphi(Y_{i,n})$

$$\mathbb{U}_n^{(T)}(\varphi) - \mathbb{E}\left(\mathbb{U}_n^{(T)}(\varphi)\right) := \mathbb{U}_{1,n}(\varphi) + \mathbb{U}_{2,n}(\varphi),$$

where

$$\mathbb{U}_{1,n}(\varphi) := \frac{1}{n} \sum_{i=1}^n \hat{H}_{1,i}(\mathbf{u}, \mathbf{x}, \varphi), \tag{97}$$

$$\mathbb{U}_{2,n}(\varphi) := \frac{(n-m)!}{(n)!} \sum_{i \in I_n^m} \xi_{i_1} \dots \xi_{i_m} H_{2,i}(z). \tag{98}$$

The convergence of $\mathbb{U}_{2,n}(\varphi)$ to zero in probability follows from Lemma 1. Hence, it is enough to show that $\mathbb{U}_{1,n}(\varphi)$ converges weakly to a Gaussian process called $\mathbb{G}(\varphi)$. In order to achieve our goal, we will go through finite-dimensional convergence and equicontinuity.

The finite-dimensional convergence simply asserts that every finite set of functions f_1, \dots, f_q in L_2 , for $\tilde{\mathbb{U}}$ the centered form of \mathbb{U} :

$$\left(\sqrt{nh^m\phi(h)} \tilde{\mathbb{U}}_{1,n}(f_1), \dots, \sqrt{nh^m\phi(h)} \tilde{\mathbb{U}}_{1,n}(f_q) \right) \tag{99}$$

converges to the corresponding finite-dimensional distributions of the process $\mathbb{G}(\varphi)$. It is sufficient to show that for every fixed collection $(a_1, \dots, a_q) \in \mathbb{R}^q$, we have

$$\sum_{j=1}^q a_j \tilde{\mathbb{U}}_{1,n}(f_j) \rightarrow N(0, v^2),$$

where

$$v^2 = \sum_{j=1}^q a_j^2 \text{Var}\left(\tilde{\mathbb{U}}_{1,n}(f_j)\right) + \sum_{s \neq r} a_s a_r \text{Cov}\left(\tilde{\mathbb{U}}_{1,n}(f_s), \tilde{\mathbb{U}}_{1,n}(f_r)\right). \tag{100}$$

Take

$$h(\cdot) = \sum_{j=1}^q a_j f_j(\cdot).$$

By linearity of $h(\cdot)$, we have to see that

$$\tilde{U}_{1,n}(h, \mathbf{i}) \rightarrow \mathbb{G}(h).$$

Let

$$N = \mathbb{E} \prod_{j=1}^m \left\{ K_2 \left(\frac{d(x_j, \mathbf{X}_{S_{ij}, A_n})}{h_n} \right) \right\}.$$

We have:

$$\begin{aligned} & \tilde{U}_{1,n}(h_n) \\ &= N^{-1} \times \frac{1}{n} \sum_{i=1}^n \frac{(n-m)!}{(n-1)!} \sum_{I_{n-1}^{m-1}(-i)} \sum_{\ell=1}^m \xi_{i_1} \dots \xi_{i_{\ell-1}} \xi_i \xi_{i_\ell} \dots \xi_{i_{m-1}} \frac{1}{\phi(h)} K_2 \left(\frac{d(x_i, \mathbf{X}_{S_{i, A_n})}}{h_n} \right) \\ & \quad \times \int h(y_1, \dots, y_{\ell-1}, Y_i, y_\ell, \dots, y_{m-1}) \prod_{\substack{j=1 \\ j \neq i}}^{m-1} \frac{1}{\phi(h)} K_2 \left(\frac{d(x_j, \mathbf{X}_{S_{ij}, A_n})}{h_n} \right) \\ & \quad \mathbb{P}(d(v_1, y_1), \dots, d(v_{\ell-1}, y_{\ell-1}), d(v_\ell, y_\ell), \dots, d(v_{m-1}, y_{m-1})), \\ & := N^{-1} \frac{1}{n} \sum_{i=1}^n \xi_i \frac{1}{\phi(h)} K_2 \left(\frac{d(x_i, \mathbf{X}_{S_{i, A_n})}}{h_n} \right) \tilde{h}(Y_i). \end{aligned}$$

The next step requires an extension of the Blocking techniques of Bernstein to the spacial process where all notions are defined in Section 9.1.

Recall that $L_n = L_{1,n} \cup L_{2,n}$ and define:

$$\mathbb{Z}_{S, A_n}(\mathbf{u}, \mathbf{x}) := \xi_i \frac{1}{\phi(h)} K_2 \left(\frac{d(x_i, \mathbf{X}_{S_{i, A_n})}}{h_n} \right) \tilde{h}(Y_i), \tag{101}$$

and

$$\mathbb{Z}_n(\boldsymbol{\ell}; \boldsymbol{\epsilon}) = \sum_{i: s_i \in \Gamma_n(\boldsymbol{\ell}; \boldsymbol{\epsilon}) \cap R_n} \mathbb{Z}_{S, A_n}(\mathbf{u}, \mathbf{x}) = \left(\mathbb{Z}_n^{(1)}(\boldsymbol{\ell}; \boldsymbol{\epsilon}), \dots, \mathbb{Z}_n^{(p)}(\boldsymbol{\ell}; \boldsymbol{\epsilon}) \right)'. \tag{102}$$

Then, we have

$$\begin{aligned} \tilde{U}_{1,n}(h_n) &= \sum_{i=1}^n \mathbb{Z}_{S, A_n}(\mathbf{u}, \mathbf{x}) \\ &= \sum_{\boldsymbol{\ell} \in L_n} \mathbb{Z}_n(\boldsymbol{\ell}; \boldsymbol{\epsilon}_0) + \underbrace{\sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} \sum_{\boldsymbol{\ell} \in L_{1,n}} \mathbb{Z}_n(\boldsymbol{\ell}; \boldsymbol{\epsilon})}_{=: \mathbb{Z}_{2,n}(\boldsymbol{\epsilon})} + \underbrace{\sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} \sum_{\boldsymbol{\ell} \in L_{2,n}} \mathbb{Z}_n(\boldsymbol{\ell}; \boldsymbol{\epsilon})}_{=: \mathbb{Z}_{3,n}(\boldsymbol{\epsilon})} \\ &=: \mathbb{Z}_{1,n} + \sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} \mathbb{Z}_{2,n}(\boldsymbol{\epsilon}) + \sum_{\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0} \mathbb{Z}_{3,n}(\boldsymbol{\epsilon}). \end{aligned} \tag{103}$$

Lemma A8 proves that $\mathbb{Z}_{2,n}$ and $\mathbb{Z}_{3,n}$, for $\boldsymbol{\epsilon} \neq \boldsymbol{\epsilon}_0$, are asymptotically negligible. Treating now the variance of $\mathbb{Z}_{1,n}$ is clear; first, mixing conditions are used to replace large blocks with independent random variables, and then Lyapunov’s condition for the central limit theorem is applied to the sum of independent random variables. Similary to the proof of Proposition 1 using Lemma A4, as in Equation (59), observe that

$$\sup_{t>0} \left| P_{\cdot|S}(\mathbb{Z}_{1,n} > t) - P_{\cdot|S}(\check{\mathbb{Z}}_{1,n} > t) \right| \leq C \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d), \tag{104}$$

where $\{\check{Z}_n(\ell; \epsilon) : \ell \in L_n\}$ denotes a sequence of independent random vectors in \mathbb{R}^p under $P_{\cdot|\mathcal{S}}$ such that

$$\check{Z}_n(\ell; \epsilon) \stackrel{d}{=} Z_n(\ell; \epsilon), \text{ under } P_{\mathcal{S}}, \ell \in L_n.$$

Applying Lyapunov’s condition for the central limit theorem for the sum of independent random variables, the remaining condition of finite-dimensional convergence must be established.

We end up with the asymptotic equicontinuity. We have to prove that:

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{P} \left\{ \sqrt{nh^m \phi(h)} \|\tilde{U}_{1,n}(h_n, \mathbf{i})\|_{\mathcal{F}\mathcal{K}(\delta, \|\cdot\|_p)} > \epsilon \right\} = 0, \tag{105}$$

where

$$\begin{aligned} \mathcal{F}\mathcal{K}(\delta, \|\cdot\|_p) &:= \{ \tilde{U}'_{1,n}(h_n) - \tilde{U}''_{1,n}(h_n) : \\ &\|\tilde{U}'_{1,n}(h_n) - \tilde{U}''_{1,n}(h_n)\| < \delta, \tilde{U}'_{1,n}(h_n), \tilde{U}''_{1,n}(h_n) \in \mathcal{F}\mathcal{K} \}, \end{aligned} \tag{106}$$

for

$$\begin{aligned} \tilde{U}'_{1,n}(h_n) &= N^{-1} \frac{1}{n} \sum_{i=1}^n \zeta_i \frac{1}{\phi(h)} K_{2,1} \left(\frac{d(x_i, X_{S_i, A_n})}{h_n} \right) \tilde{h}_1(Y_i) - \mathbb{E}(\mathbb{U}'_{1,n}(h_n)) \\ \tilde{U}''_{1,n}(h_n) &= N^{-1} \frac{1}{n} \sum_{i=1}^n \zeta_i \frac{1}{\phi(h)} K_{2,2} \left(\frac{d(x_i, X_{S_i, A_n})}{h_n} \right) \tilde{h}_2(Y_i) - \mathbb{E}(\mathbb{U}''_{1,n}(h_n)) \end{aligned} \tag{107}$$

At this point, we will adapt the chaining technique found in [82] and use it for the conditional setting with the locally stationary process in [152] but for random fields, as in Lemma 1.

Using the same strategy also as in Lemma 1 to pass from the sequence of locally stationary random variables to the stationary one and find that, for $\zeta_i = (\eta_i, \varsigma_i)$, the independent blocks sequences:

$$\begin{aligned} &\mathbb{P} \left\{ \left\| (n\phi(h))^{-1/2} h^{m/2} N^{-1} \sum_{i=1}^n \left(\zeta_i K_2 \left(\frac{d(x_i, X_i)}{h} \right) \tilde{h}(Y_i) - \mathbb{E}(\mathbb{U}_{1,n}(h_n)) \right) \right\|_{\mathcal{F}\mathcal{K}(b, \|\cdot\|_p)} > \epsilon \right\} \\ &\leq 2\mathbb{P} \left\{ \left\| (n\phi(h))^{-1/2} h^{m/2} N^{-1} \sum_{\ell \in L_n} \sum_{i: S_i \in \Gamma_n(\ell; \epsilon_0) \cap R_n} \left(\zeta_i K_2 \left(\frac{d(x_i, \eta_i)}{h} \right) \tilde{h}(\varsigma_i) \right. \right. \\ &\quad \left. \left. - \mathbb{E}(\mathbb{U}_{1,n}(h_n)) \right) \right\|_{\mathcal{F}\mathcal{K}(b, \|\cdot\|_p)} > \epsilon' \right\} + C \left(\frac{A_n}{A_{1,n}} \right)^d \beta(A_{2,n}; A_n^d) + o_{\mathbb{P}}(1). \end{aligned} \tag{108}$$

Taking advantage of the condition (E2) in Assumption 6, we obtain $\beta(A_{2,n}; A_n^d) \rightarrow 0$ as $n \rightarrow \infty$; then, it is simply a matter of placing the first phrase in the right-hand side of (109). Due to the fact that the blocks are independent, we symmetricize 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 is important to notice that the sequence $\{\epsilon_j\}_{j \in \mathbb{N}^*}$ is independent of the sequence $\{\xi_i = (\zeta_i, \zeta_i)\}_{i \in \mathbb{N}^*}$; therefore, it remains to establish, for all $\epsilon > 0$ and $\delta \rightarrow 0$,

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{P} \left\{ \left\| (n\phi(h))^{-1/2} h^{m/2} N^{-1} \sum_{\ell \in L_n} \sum_{i: s_i \in \Gamma_n(\ell; \epsilon_0) \cap R_n} \left(\xi_i K_2 \left(\frac{d(x_i, \eta_i)}{h} \right) \tilde{h}(\zeta_i) - \mathbb{E}(\mathbb{U}_{1,n}(h_n, \mathbf{i})) \right) \right\|_{\mathcal{F}_{\mathcal{H}}(\|\cdot\|_p)} > \epsilon \right\} < \delta. \tag{109}$$

Define the semi-norm:

$$\begin{aligned} \tilde{d}_{n\phi,2} := & \left((n\phi(h))^{-1/2} h^{m/2} N^{-1} \sum_{\ell \in L_n} \sum_{i: s_i \in \Gamma_n(\ell; \epsilon_0) \cap R_n} \left| \left(\xi_i K_{2,1} \left(\frac{d(x_i, \eta_i)}{h} \right) \tilde{h}_1(\zeta_i) - \mathbb{E}(\mathbb{U}'_{1,n}(h_n, \mathbf{i})) \right) - \left(\xi_i K_{2,2} \left(\frac{d(x_i, \eta_i)}{h} \right) \tilde{h}_2(\zeta_i) - \mathbb{E}(\mathbb{U}''_{1,n}(h_n, \mathbf{i})) \right) \right|^2 \right)^{1/2} \end{aligned} \tag{110}$$

and the covering number defined for any class of functions \mathcal{E} by:

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

Because of the latter, we are able to bound (109) (more details are in [83]). In the same way, as in [83] and before in [82], as a result of the independence between the blocks and Assumption 7 (C3), and by applying ([151], Lemma 5.2), the equicontinuity is achieved, and then the weak convergence is achieved, too.

Now, we need to show that:

$$\mathbb{P} \left\{ \left\| \mu_n^{(R)}(\varphi, \mathbf{t}) \right\|_{\mathcal{F}_m \mathcal{X}^m} > \lambda \right\} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

For clarity purposes, we restrict ourselves to $m = 2$. Using the same notation as in Lemma 1, we have the following decomposition:

$$\begin{aligned} \mu_n^{(R)}(\varphi, \mathbf{i}) &= \sqrt{nh^{m+d}\phi(h_n)} \left\{ \mathbb{U}_n^{(R)}(\varphi, \mathbf{i}) - \mathbb{E} \left(\mathbb{U}_n^{(R)}(\varphi, \mathbf{i}) \right) \right\} \\ &= \frac{\sqrt{nh^{m+d}\phi(h_n)}}{n(n-1)} \sum_{i_1 \neq i_2}^n \xi_{i_1} \xi_{i_2} \left\{ G_{\varphi, \mathbf{t}}^{(R)} \left(((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2})) \right) - \mathbb{E} \left[G_{\varphi, \mathbf{i}}^{(R)} \left(((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2})) \right) \right] \right\} \\ &\leq \frac{1}{\sqrt{nh^{m+d}\phi(h_n)}} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon_0) \cap R_n} \phi(h_n) \xi_{i_1} \xi_{i_2} \left\{ G_{\varphi, \mathbf{i}}^{(R)} \left((X_i, X_j), (Y_i, Y_j) \right) - \mathbb{E} \left[G_{\varphi, \mathbf{i}}^{(R)} \left((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}) \right) \right] \right\} \\ &+ \frac{1}{\sqrt{nh^{m+d}\phi(h_n)}} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\substack{i_2: s_{i_2} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n \\ i_1 \neq i_2}} \phi(h_n) \xi_{i_1} \xi_{i_2} \left\{ G_{\varphi, \mathbf{i}}^{(R)} \left((X_i, X_j), (Y_i, Y_j) \right) - \mathbb{E} \left[G_{\varphi, \mathbf{i}}^{(R)} \left((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}) \right) \right] \right\} \\ &+ 2 \frac{1}{\sqrt{nh^{m+d}\phi(h_n)}} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \end{aligned}$$

$$\begin{aligned}
 & \phi(h_n)\xi_{i_1}\xi_{i_2}\left\{G_{\varphi,i}^{(R)}((X_i, X_j), (Y_i, Y_j)) - \mathbb{E}\left[G_{\varphi,i}^{(R)}((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}))\right]\right\} \\
 +2 & \frac{1}{\sqrt{nh^{m+d}}\phi(h_n)} \sum_{\ell_1 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\Delta_1} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \\
 & \phi(h_n)\xi_{i_1}\xi_{i_2}\left\{G_{\varphi,t}^{(R)}((X_i, X_j), (Y_i, Y_j)) - \mathbb{E}\left[G_{\varphi,t}^{(R)}((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}))\right]\right\} \\
 + & \frac{1}{\sqrt{nh^{m+d}}\phi(h_n)} \sum_{\substack{\ell_1 \neq \ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \\
 & \phi(h_n)\xi_{i_1}\xi_{i_2}\left\{G_{\varphi,t}^{(R)}((X_i, X_j), (Y_i, Y_j)) - \mathbb{E}\left[G_{\varphi,t}^{(R)}((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}))\right]\right\} \\
 + & \frac{1}{\sqrt{nh^{m+d}}\phi(h_n)} \sum_{\substack{\ell_1 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_1 < i_2 : s_{i_1}, s_{i_2} \in \Gamma_n(\ell_1;\epsilon) \cap R_n} \\
 & \phi(h_n)\xi_{i_1}\xi_{i_2}\left\{G_{\varphi,t}^{(R)}((X_i, X_j), (Y_i, Y_j)) - \mathbb{E}\left[G_{\varphi,t}^{(R)}((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}))\right]\right\} \\
 =: & \text{I}' + \text{II}' + \text{III}' + \text{IV}' + \text{V}' + \text{VI}' .
 \end{aligned}$$

We shall employ blocking arguments and evaluate the terms that result. We begin by examining the first I'. We obtain

$$\begin{aligned}
 & \mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n)\xi_{i_1}\xi_{i_2}\left\{G_{\varphi,i}^{(R)}((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}))\right.\right.\right. \\
 & \quad \left.\left.\left.- \mathbb{E}\left[G_{\varphi,t}^{(R)}((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}))\right]\right\}\right\|_{\mathcal{F}_2, \mathcal{H}^2} > \delta\right\} \\
 \leq & \mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n)\xi_{i_1}\xi_{i_2}\right.\right. \\
 & \quad \left.\left.\left\{G_{\varphi,i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2})) - \mathbb{E}\left[G_{\varphi,i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2}))\right]\right\}\right\|_{\mathcal{F}_2, \mathcal{H}^2} > \delta\right\} \\
 & + C\left(\frac{A_n}{A_{1,n}}\right)^d \beta(A_{2,n}; A_n^d).
 \end{aligned}$$

Recall that for all $\varphi \in \mathcal{F}_m$, and:

$$x \in \mathcal{H}^2, \mathbf{y} \in \mathcal{Y}^2 : \mathbf{1}_{\{d(x, X_{i,n}) \leq h\}} F(\mathbf{y}) \geq \varphi(\mathbf{y}) K_2 \left(\frac{d(x_i, X_{s_i, A_n})}{h_n}\right).$$

Hence, by the symmetry of $F(\cdot)$:

$$\begin{aligned}
 & \left\|\frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n)\xi_{i_1}\xi_{i_2}\left\{G_{\varphi,i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2}))\right.\right. \\
 & \quad \left.\left.- \mathbb{E}\left[G_{\varphi,t}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2}))\right]\right\}\right\|_{\mathcal{F}_2, \mathcal{H}^2} \\
 \lesssim & \left|\frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n)\xi_{i_1}\xi_{i_2}\left\{F(\zeta_i, \zeta_j) \mathbf{1}_{\{F > \lambda_n\}}\right.\right. \tag{111}
 \end{aligned}$$

$$-\mathbb{E}\left[F(\zeta_i, \zeta_j) \mathbf{1}_{\{F>\lambda_n\}}\right]\Bigg|.$$

We are going to use Chebyshev’s inequality, Hoeffding’s trick and inequality, respectively, to obtain:

$$\begin{aligned} & \mathbb{P}\left\{\left|\frac{1}{\sqrt{n\phi(h)}} \sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} \left\{F(\zeta_i, \zeta_j) \mathbf{1}_{\{F>\lambda_n\}}\right\} \right. \right. \\ & \left. \left. - \mathbb{E}\left[F(\zeta_i, \zeta_j) \mathbf{1}_{\{F>\lambda_n\}}\right]\right\} > \delta\right\} \\ & \lesssim \delta^{-2} n^{-1} \phi^{-1}(h) \text{Var}\left(\sum_{\ell_1 \neq \ell_2 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} F(\zeta_i, \zeta_j) \mathbf{1}_{\{F>\lambda_n\}}\right) \tag{112} \\ & \lesssim c_2 \llbracket L_n \rrbracket \delta^{-2} n^{-1} \phi^{-1}(h) \text{Var}\left(\sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} F(\zeta_i, \zeta_j) \mathbf{1}_{\{F>\lambda_n\}}\right) \\ & \lesssim 2c_2 \llbracket L_n \rrbracket \delta^{-2} n^{-2} \phi^{-1}(h) \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} \mathbb{E}\left[(F(\zeta_1, \zeta_2))^2 \mathbf{1}_{\{F>\lambda_n\}}\right]. \end{aligned}$$

Under Assumption 7 (iii), we have for each $\lambda > 0$:

$$\begin{aligned} & c_2 \llbracket L_n \rrbracket \delta^{-2} n^{-2} \phi^{-1}(h_n) \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} \mathbb{E}\left[(F(\zeta_1, \zeta_2))^2 \mathbf{1}_{\{F>\lambda_n\}}\right] \\ & = c_2 \llbracket L_n \rrbracket \delta^{-2} n^{-2} \phi^{-1}(h_n) \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} \\ & \times \int_0^\infty \mathbb{P}\left\{(F(\zeta_1, \zeta_2))^2 \mathbf{1}_{\{F>\lambda_n\}} \geq t\right\} dt \\ & = c_2 \llbracket L_n \rrbracket \delta^{-2} n^{-2} \phi^{-1}(h_n) \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} \int_0^{\lambda_n} \mathbb{P}\{F > \lambda_n\} dt \\ & + c_2 \llbracket L_n \rrbracket \delta^{-2} n^{-2} \phi^{-1}(h_n) \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon_0) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} \int_{\lambda_n}^\infty \mathbb{P}\{(F)^2 > t\} dt, \end{aligned}$$

converging to 0 as $n \rightarrow \infty$. Terms II’, V’ and VI’ will be handled the same way as the last term was. The terms II’, VI’ do not follow the same line because the variables $\{\zeta_i, \zeta_j\}_{\epsilon=\epsilon_0}$ (or $\{\zeta_i, \zeta_j\}_{\epsilon \neq \epsilon_0}$ for VI’) belong to the same blocks. Term IV’ can be deduced from the study of Terms I’ and III’. Considering the term III’, we have

$$\mathbb{P}\left\{\left\|\frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \in L_n} \sum_{i_1:s_{i_1} \in \Gamma_n(\ell_1;\epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2:s_{i_2} \in \Gamma_n(\ell_2;\epsilon) \cap R_n} \phi(h_n) \zeta_{i_1} \zeta_{i_2} \right. \right. \\ \left. \left. \left\{G_{\varphi,i}^{(R)}((X_i, X_j), (Y_i, Y_j)) - \mathbb{E}\left[G_{\varphi,i}^{(R)}((X_{i_1}, X_{i_2}), (Y_{i_1}, Y_{i_2}))\right]\right\}\right\|_{\mathcal{F}_2 \mathcal{K}^2} > \delta\right\}$$

$$\begin{aligned} &\leq \mathbb{P} \left\{ \left\| \frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \phi(h_n) \xi_{i_1} \xi_{i_2} \right. \right. \\ &\quad \left. \left. \left\{ G_{\varphi, i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2})) - \mathbb{E} \left[G_{\varphi, i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2})) \right] \right\} \right\|_{\mathcal{F}_2, \mathcal{H}^2} > \delta \right\} \\ &+ \frac{\llbracket L_n \rrbracket A_{1,n}^d A_{2,n}^d \beta(A_{2,n}; A_n^d)}{\sqrt{n\phi(h_n)}}. \end{aligned} \tag{113}$$

We have also

$$\begin{aligned} &\mathbb{P} \left\{ \left\| \frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \phi(h_n) \xi_{i_1} \xi_{i_2} \right. \right. \\ &\quad \left. \left. \left\{ G_{\varphi, i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2})) - \mathbb{E} \left[G_{\varphi, i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2})) \right] \right\} \right\|_{\mathcal{F}_2, \mathcal{H}^2} > \delta \right\} \\ &\leq \mathbb{P} \left\{ \left\| \frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \phi(h_n) \xi_{i_1} \xi_{i_2} \right. \right. \\ &\quad \left. \left. \left\{ G_{\varphi, i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2})) - \mathbb{E} \left[G_{\varphi, i}^{(R)}((\varsigma_{i_1}, \varsigma_{i_2}), (\zeta_{i_1}, \zeta_{i_2})) \right] \right\} \right\|_{\mathcal{F}_2, \mathcal{H}^2} > \delta \right\}. \end{aligned}$$

Since (111) is still true, the problem can be reduced to

$$\begin{aligned} &\mathbb{P} \left\{ \left\| \frac{1}{\sqrt{n\phi(h_n)}} \sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \phi(h_n) \xi_{i_1} \xi_{i_2} \right. \right. \\ &\quad \left. \left. \left\{ F(\zeta_i, \zeta_j) \mathbf{1}_{\{F > \lambda_n\}} - \mathbb{E} \left[F(\zeta_i, \zeta_j) \mathbf{1}_{\{F > \lambda_n\}} \right] \right\} \right\| > \delta \right\} \\ &\lesssim \delta^{-2} n^{-1} \phi(h_n) \text{Var} \left(\sum_{\ell_1 \in L_n} \sum_{i_1: s_{i_1} \in \Gamma_n(\ell_1; \epsilon_0) \cap R_n} \sum_{\Delta_2} \sum_{\substack{\ell_2 \in L_{1,n} \cup L_{2,n} \\ \epsilon \neq \epsilon_0}} \sum_{i_2: s_{i_2} \in \Gamma_n(\ell_2; \epsilon) \cap R_n} \phi(h_n) \xi_{i_1} \xi_{i_2} \right. \\ &\quad \left. \times F(\zeta_i, \zeta_j) \mathbf{1}_{\{F > \lambda_n\}} \right), \end{aligned}$$

the identical technique is followed as in (112). The remainder has just been demonstrated to be asymptotically negligible. Finally, with $\hat{r}^{(m)}(\varphi, \mathbf{x}, \mathbf{u}) \rightarrow \mathbb{E}(\mathbb{U}_n(\varphi, \mathbf{i}))$, and for $(\mathbb{U}_n(1, \mathbf{i})) \xrightarrow{\mathbb{P}} 1$, the weak convergence of our estimator is accomplished. \square

Author Contributions: I.S. and S.B.: conceptualization, methodology, investigation, writing—original draft, writing—review and editing. All authors contributed equally to the writing of this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the Special Issue Editor of the Special Issue on “Current Developments in Theoretical and Applied Statistics”, Christophe Chesneau for the invitation. We extend our sincere thanks to the Editor-in-Chief, Associate Editor and several referees for their constructive comments that lead to numerous improvements over a previous version.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This appendix contains supplementary information that is an essential part of providing a more comprehensive understanding of the paper.

Assumption A1.

(KD1) (KB2) in Assumption 2 holds.

(KD2) For any $\alpha \in \mathbb{Z}^d$ where $|\alpha| = 1, 2, \partial^\alpha f_S(\mathbf{s})$ exists and is continuous on $(0, 1)^d$.

Define

$$\widehat{f}_S(\mathbf{u}) = \frac{1}{nh^d} \sum_{j=1}^n \bar{K}_h(\mathbf{u} - \mathbf{S}_{0,j}).$$

Lemma A1 ([153], Theorem 2). Under Assumption A1 and $h \rightarrow 0$ such that $nh^d / (\log n) \rightarrow \infty$ as $n \rightarrow \infty$, we have that

$$\sup_{\mathbf{u} \in [0,1]^d} |\widehat{f}_S(\mathbf{u}) - f_S(\mathbf{u})| = O\left(\sqrt{\frac{\log n}{nh^d}} + h^2\right) P_S - a.s.$$

Lemma A2. Let $I_h = [C_1h, 1 - C_1h]$. Suppose that kernel K_1 satisfies Assumption A1 part(i). Then for $q = 0, 1, 2$ and $m > 1$:

$$\begin{aligned} \sup_{\mathbf{u} \in I_h} & \left| \frac{1}{n^m h^{md}} \sum_{\mathbf{i} \in I_h^m} \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{S}_{0,i_j}}{h_n}\right) \left(\frac{\mathbf{u}_j - \mathbf{S}_{0,i_j}}{h}\right)^q \right. \\ & \left. - \int_{\mathbb{R}^{md}} \frac{1}{h^{md}} \prod_{j=1}^m \left\{ \bar{K}\left(\frac{\mathbf{u}_j - \boldsymbol{\omega}_j}{h_n}\right) \left(\frac{\mathbf{u}_j - \boldsymbol{\omega}_j}{h}\right)^q \right\} f_S(\boldsymbol{\omega}_j) \prod_{j=1}^m d\boldsymbol{\omega}_j \right| = O\left(\sqrt{\frac{\log n}{nh^{dm}}}\right) P_S - a.s. \end{aligned}$$

Lemma A3. Suppose that kernel \bar{K} satisfies Assumption A1. Let $g : [0, 1]^{md} \times \mathcal{H}^m \rightarrow \mathbb{R}$, $(\mathbf{u}, \mathbf{x}) \mapsto g(\mathbf{u}, \mathbf{x})$ be a function continuously partially differentiable with respect to \mathbf{u}_j . For $k = 1, 2$, we have

$$\begin{aligned} \sup_{\mathbf{u} \in I_h} & \left| \frac{1}{n^m h^{md}} \sum_{\mathbf{i} \in I_h^m} \prod_{j=1}^m \bar{K}\left(\frac{\mathbf{u}_j - \mathbf{S}_{0,i_j}}{h_n}\right)^k g(\mathbf{S}_{0,i_j}, \mathbf{x}_j) - \prod_{j=1}^m \kappa_k f_S(\mathbf{u}_j) g(\mathbf{u}_j, \mathbf{x}_j) \right| \\ & = O\left(\sqrt{\frac{\log n}{nh^{md}}}\right) + o(h), P_S - a.s. \end{aligned} \tag{A1}$$

where

$$\kappa_k = \int_{\mathbb{R}^d} \bar{K}^k(\mathbf{x}) d\mathbf{x}.$$

For any probability measure Q on a product measure space $(\Omega_1 \times \Omega_2, \Sigma_1 \times \Sigma_2)$, we may define the β -mixing coefficients as follows:

Definition A1 ([149], Definition 2.5). Let Q_1 and Q_2 be the marginal probability measures of Q on (Ω_1, Σ_1) and (Ω_2, Σ_2) , respectively. We set

$$\beta(\Sigma_1, \Sigma_2, Q) = P \sup\{|Q(B | \Sigma_1) - Q_2(B)| : B \in \Sigma_2\}.$$

The following lemma holds true for every finite n and is essential for the generation of independent blocks for β -mixing sequences.

Lemma A4 ([149], Corollary 2.7). Let $m \in \mathbb{N}$ and let Q denote a probability measure on a product space $(\prod_{i=1}^m \Omega_i, \prod_{i=1}^m \Sigma_i)$ with the associated marginal measures Q_i on (Ω_i, Σ_i) . Assume that h is a bounded measurable function on the product probability space in such a way that $|h| \leq M_h < \infty$. For $1 \leq a \leq b \leq m$, let Q_a^b be the marginal measure on $(\prod_{i=a}^b \Omega_i, \prod_{i=a}^b \Sigma_i)$. For a given $\tau > 0$, suppose that, for all $1 \leq k \leq m - 1$,

$$\|Q - Q_1^k \times Q_{k+1}^m\|_{TV} \leq 2\tau \tag{A2}$$

where $Q_1^k \times Q_{k+1}^m$ is the product measure and $\|\cdot\|_{TV}$ is the total variation. Then

$$|Qh - Ph| \leq 2M_h(m - 1)\tau,$$

where $P = \prod_{i=1}^m Q_i, Qh = \int hdQ$ and $Ph = \int hdP$.

Lemma A5. Let

$$\mathcal{I}_n = \{i \in \mathbb{Z}^d : (i + (0, 1]^d) \cap R_n \neq \emptyset\}.$$

Then, we have

$$P_S \left(\sum_{j=1}^n \mathbf{1}\{A_n S_{0,j} \in (i + (0, 1]^d) \cap R_n\} > 2(\log n + nA_n^{-d}) \text{ for some } i \in \mathcal{I}_n, \text{ i.o.} \right) = 0$$

and

$$P_S \left(\sum_{j=1}^n \mathbf{1}\{A_n S_{0,j} \in \Gamma_n(\ell; \epsilon)\} > CA_{1,n}^{q(\epsilon)} A_{2,n}^{d-q(\epsilon)} nA_n^{-d} \text{ for some } \ell \in L_{1,n}, \text{ i.o.} \right) = 0$$

for any $\epsilon \in \{1, 2\}^d$, where ‘‘i.o.’’ stands for infinitely often.

Proof. See the proof in ([93], Lemma A.1) for each statement. \square

Remark A1. Lemma A5 implies that each $\Gamma_n(\ell; \epsilon)$ contains at most $CA_{1,n}^{q(\epsilon)} A_{2,n}^{d-q(\epsilon)} nA_n^{-d}$ samples P_S -almost surely.

Lemma A6. Under Assumptions 2 and 3, Condition (B1) in 4–6 and A1, we have:

$$E_{\cdot|S} \left[(\bar{S}_n(\ell; \epsilon))^2 \right] \leq CA_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{md} \phi(h).$$

Appendix A.1. Proof of Lemma A6

We have

$$E_{\cdot|S} \left[(\bar{S}_n(\ell; \epsilon))^2 \right]$$

$$= \sum_{i:s_i \in \Gamma_n(\ell; \epsilon) \cap R_n} E_{\cdot|S} \left[\bar{S}_{s_i, A_n}^2(\mathbf{u}, \mathbf{x}) \right] + \sum_{i \neq j: s_i, s_j \in \Gamma_n(\ell; \epsilon) \cap R_n} E_{\cdot|S} \left[\bar{S}_{s_i, A_n}(\mathbf{u}, \mathbf{x}) \bar{S}_{s_j, A_n}(\mathbf{u}, \mathbf{x}) \right]$$

where

$$\begin{aligned} & \sum_{i:s_i \in \Gamma_n(\ell; \epsilon) \cap R_n} E_{\cdot|S} \left[\bar{S}_{s_i, A_n}^2(\mathbf{u}, \mathbf{x}) \right] \\ & \leq (n-1)^{-2m} \sum_{\substack{I_{n-1}^{m-1}(-i) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{i_1}^2 \dots \tilde{\zeta}_{i_{\ell-1}}^2 \tilde{\zeta}_{i_\ell}^2 \dots \tilde{\zeta}_{i_{m-1}}^2 \left(\int W_{s_{(1, \dots, \ell-1, \ell, \dots, m-1)}, A_n} \prod_{\substack{j=1 \\ j \neq i}}^{m-1} \frac{1}{\phi(h)} \right. \\ & \quad \left. K_2 \left(\frac{d(x_j, \nu_{s_j, A_n})}{h} \right) \mathbb{P}(dv_1, \dots, dv_{\ell-1}, dv_\ell, \dots, dv_{m-1}) \right)^2 \\ & \quad \left\{ E_{\cdot|S} \left(\frac{1}{\phi^2(h)} K_2^2 \left(\frac{d(x_i, X_{s_i, A_n})}{h} \right) W_{s_i, A_n}^2 \right) \right. \\ & \quad \left. + \left[E_{\cdot|S} \left(\frac{1}{\phi^2(h)} K_2 \left(\frac{d(x_i, X_{s_i, A_n})}{h} \right) W_{s_i, A_n} \right) \right]^2 \right\} \tag{A3} \\ & \leq C \phi^2(h) (n-1)^{-2m} \sum_{\substack{I_{n-1}^{m-1}(-i) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{i_1}^2 \dots \tilde{\zeta}_{i_{\ell-1}}^2 \tilde{\zeta}_{i_\ell}^2 \dots \tilde{\zeta}_{i_{m-1}}^2 \quad P_S - a.s. \end{aligned}$$

Likewise, we can see that

$$\begin{aligned} & E_{\cdot|S} \left[\bar{S}_{s_i, A_n}(\mathbf{u}, \mathbf{x}) \bar{S}_{s_j, A_n}(\mathbf{u}, \mathbf{x}) \right] \\ & \leq C \phi^2(h) (n-1)^{-2m} \sum_{\substack{I_{n-1}^{m-1}(-i) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{i_1}^2 \dots \tilde{\zeta}_{i_{\ell-1}}^2 \tilde{\zeta}_{i_\ell}^2 \dots \tilde{\zeta}_{i_{m-1}}^2 \quad P_S - a.s. \tag{A4} \end{aligned}$$

Applying Lemma A5 and Lemma A1 to find that

$$\begin{aligned} & \sum_{i:s_i \in \Gamma_n(\ell; \epsilon) \cap R_n} \bar{K}_h^2 \left(u - \frac{s_j}{A_n} \right) \times (n-1)^{-2m} \sum_{\substack{I_{n-1}^{m-1}(-i) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{i_1}^2 \dots \tilde{\zeta}_{i_{\ell-1}}^2 \tilde{\zeta}_{i_\ell}^2 \dots \tilde{\zeta}_{i_{m-1}}^2 \\ & \leq C \sum_{i:s_i \in \Gamma_n(\ell; \epsilon) \cap R_n} \bar{K}_h \left(u - \frac{s_i}{A_n} \right) \times (n-1)^{-2m} \sum_{\substack{I_{n-1}^{m-1}(-i) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{i_1} \dots \tilde{\zeta}_{i_{\ell-1}} \tilde{\zeta}_{i_\ell} \dots \tilde{\zeta}_{i_{m-1}} \\ & \leq Ch^{md} \mathbb{I} \{ i : s_i \in \Gamma_n(\ell; \epsilon) \cap R_n \} \\ & \leq Ch^{md} A_{1,n}^{d-1} A_{2,n} (nA^{-d} + \log n), P_S - a.s., \end{aligned}$$

and

$$\begin{aligned} & \sum_{i \neq j: s_i, s_j \in \Gamma_n(\ell; \epsilon) \cap R_n} \bar{K}_h \left(u - \frac{s_i}{A_n} \right) \bar{K}_h \left(u - \frac{s_j}{A_n} \right) \\ & \quad \times (n-1)^{-m} \left(\sum_{\substack{I_{n-1}^{m-1}(-i) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{i_1} \dots \tilde{\zeta}_{i_{\ell-1}} \tilde{\zeta}_{i_\ell} \dots \tilde{\zeta}_{i_{m-1}} \sum_{\substack{I_{n-1}^{m-1}(-j) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{j_1} \dots \tilde{\zeta}_{j_{\ell-1}} \tilde{\zeta}_{j_\ell} \dots \tilde{\zeta}_{j_{m-1}} \right)^2 \\ & \leq \left(\sum_{j:s_j \in \Gamma_n(\ell; \epsilon) \cap R_n} \bar{K}_h \left(u - \frac{s_j}{A_n} \right) (n-1)^{-m} \sum_{\substack{I_{n-1}^{m-1}(-j) \\ \ell=1}}^m \sum_{\ell=1}^m \tilde{\zeta}_{j_1} \dots \tilde{\zeta}_{j_{\ell-1}} \tilde{\zeta}_{j_\ell} \dots \tilde{\zeta}_{j_{m-1}} \right)^2 \\ & \leq Ch^{2md} \mathbb{I} \{ j : s_j \in \Gamma_n(\ell; \epsilon) \cap R_n \}^2 \end{aligned}$$

$$\leq Ch^{2md} A_{1,n}^{2(d-1)} A_{2,n}^2 (nA^{-d} + \log n)^2, P_S - a.s.$$

Since

$$A_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{md} \phi(h) \leq A_{1,n}^d (nA_n^{-d} + \log n) h^{md} \phi(h) = o(1),$$

we have

$$\begin{aligned} E_{\cdot|S} [(\bar{S}_n(\ell; \epsilon))^2] &\leq C \left\{ A_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{md} \phi(h) \right. \\ &\quad \left. + A_{1,n}^{2(d-1)} A_{2,n}^2 (n^2 A_n^{-2d} + \log^2 n) h^{2(d)} \phi^2(h) \right\} \\ &\leq CA_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{md} \phi(h), P_S - a.s. \end{aligned}$$

Lemma A7 (Bernstein’s inequality). Let X_1, \dots, X_n be zero-mean independent random variables. Assume that

$$\max_{1 \leq i \leq n} |X_i| \leq M < \infty, \quad a.s.$$

For all $t > 0$, we have

$$\mathbb{P} \left(\sum_{i=1}^n X_i \geq t \right) \leq \exp \left(- \frac{\frac{t^2}{2}}{\sum_{i=1}^n \mathbb{E} [X_i^2] + \frac{Mt}{3}} \right)$$

Lemma A8. Under Assumptions 2–4 and 6, we have

$$\frac{1}{nh^{md} \phi^m(h)} \text{Var}_{\cdot|S} \left(\sum_{\ell \in L_{1,n}} \mathbb{Z}_n(\ell; \epsilon) \right) = o(1), P_S - a.s. \tag{A5}$$

$$\frac{1}{nh^{md} \phi^m(h)} \text{Var}_{\cdot|S} \left(\sum_{\ell \in L_{2,n}} \mathbb{Z}_n(\ell; \epsilon) \right) = o(1), P_S - a.s. \tag{A6}$$

□

Appendix A.2. Proof of Lemma A8

We have

$$\begin{aligned} \frac{1}{nh^{md} \phi(h)} \text{Var}_{\cdot|S} \left(\sum_{\ell \in L_{1,n}} \mathbb{Z}_n(\ell; \epsilon) \right) &= \frac{1}{nh^{md} \phi(h)} \sum_{\ell \in L_{1,n}} E_{\cdot|S} [(\mathbb{Z}_n(\ell; \epsilon))^2] \\ &\quad + \frac{1}{nh^{md} \phi(h)} \sum_{\ell_1 \neq \ell_2 \in L_{1,n}} E_{\cdot|S} [\mathbb{Z}_n(\ell_1; \epsilon) \mathbb{Z}_n(\ell_2; \epsilon)] \\ &:= I_1 + I_2. \end{aligned} \tag{A7}$$

Using Lemma A6 and Assumption 4, it is easy to see that

$$\begin{aligned} I_1 &\leq C \frac{1}{nh^{md} \phi(h)} \left(\frac{A_n}{A_{1,n}} \right)^d A_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{md} \phi(h) \\ &\leq C \frac{A_{2,n}}{A_{1,n}} (\log n) = o(1). \end{aligned} \tag{A8}$$

For I_2 , using [Theorem 1.1] from [154], we have:

$$\begin{aligned} & E_{\cdot|S}[\mathbb{Z}_n(\ell_1; \epsilon)\mathbb{Z}_n(\ell_2; \epsilon)] \\ & \leq E_{\cdot|S} \left[|\mathbb{Z}_n(\ell_1; \epsilon)|^3 \right]^{1/3} E_{\cdot|S} \left[|\mathbb{Z}_n(\ell_2; \epsilon)|^3 \right]^{1/3} \beta^{1/3}(d(\ell_1, \ell_2)A_{2,n}, A_{1,n}^{md}) \\ & \leq E_{\cdot|S} \left[|\mathbb{Z}_n(\ell_1; \epsilon)|^3 \right]^{1/3} E_{\cdot|S} \left[|\mathbb{Z}_n(\ell_2; \epsilon)|^3 \right]^{1/3} \beta_1^{1/3}(d(\ell_1, \ell_2)A_{2,n})g_1^{1/3}(A_{1,n}^{md}). \end{aligned}$$

The first inequality holds using Equation (8), and for $d(\ell_1, \ell_2) = \max_{1 \leq j \leq d} |\ell_1^j - \ell_2^j|$. Using the same strategy as Lemma A6, we have

$$E_{\cdot|S} \left[|\mathbb{Z}_n(\ell_1; \epsilon)|^3 \right] \leq CA_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{md},$$

and

$$E_{\cdot|S} \left[|\mathbb{Z}_n(\ell_2; \epsilon)|^3 \right] \leq CA_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{md}.$$

Note that for $\ell_1, \ell_2 \in L_{1,n}, \Gamma(\ell_1; \epsilon_0)$ and $\Gamma(\ell_2; \epsilon_0)$ in R_n are separated by the $(\ell^1 -)$ distance

$$d(\Gamma(\ell_1; \epsilon_0), \Gamma(\ell_2; \epsilon_0)) \geq ((|\ell_1 - \ell_2| - d) + A_{3,n}) + A_{2,n}.$$

$$\begin{aligned} I_2 & \leq C \frac{\left(A_{1,n}^{d-1} A_{2,n} (nA_n^{-d} + \log n) h^{p+d} \right)^{2/3}}{nh^{d+p}} \\ & \quad \times \sum_{\ell_1, \ell_2 \in L_{1,n}, \ell_1 \neq \ell_2} \beta_1^{1/3}((|\ell_1 - \ell_2| - d) + A_{3,n} + A_{2,n}) g_1^{1/3}(A_{1,n}^d) \tag{A9} \\ & \leq C \left\{ \left(\frac{1}{nh^{d+p}} \right)^{1/3} \left(\frac{A_{1,n}}{A_n} \right)^{2d/3} \left(\frac{A_{2,n}}{A_{1,n}} \right)^{2/3} + \frac{A_{1,n}^{(d-1)/3} A_{2,n}^{1/3} (\log n)^{1/3}}{nh^{(d+p)/3}} \right\} \\ & \quad \times g_1^{1/3}(A_{1,n}^d) \left\{ \beta_1^{1/3}(A_{2,n}) + \sum_{k=1}^{A_n/A_{1,n}} k^{d-1} \beta_1^{1/3}(kA_{3,n} + A_{2,n}) \right\} = o(1) \end{aligned}$$

The last inequality follows using Assumption 4 and for

$$|\ell_1 - \ell_2| = \sum_{j=1}^d |\ell_{1,j} - \ell_{2,j}|.$$

Equation (A6) could be treated similarly to (A5).

Remark A2. In order to prove that the summation over the small block is asymptotically negligible, we can use the method of [87] where they used to pass from the dependence structure of variables to the independence as a first step; then, they proved the convergence of second-order expectation to zero using a maximal inequality. This method avoids the treatment of covariance, and it is based on the use of maximal inequality.

Proposition A1 ([27], Proposition 3.6). Let $\{X_i : i \in n\}$ be a process satisfying, for $m \geq 1$:

$$\left(\mathbb{E} \|X_i - X_j\|^p \right)^{1/p} \leq \left(\frac{p-1}{q-1} \right)^{m/2} \left(\mathbb{E} \|X_i - X_j\|^q \right)^{1/q}, \quad 1 < q < p < \infty,$$

and the semi-metric:

$$\rho(j, i) = \left(\mathbb{E} \|X_i - X_j\|^2 \right)^{1/2}.$$

There exists a constant $K = K(m)$ such that:

$$\mathbb{E} \sup_{i,j \in n} \|X_i - X_j\| \leq K \int_0^D [\log N(\epsilon, n, \rho)]^{m/2} d\epsilon,$$

where D is the ρ -diameter of n .

Lemma A9 ([155]). Let X_1, \dots, X_n be a sequence of independent random elements taking values in a Banach space $(B, \|\cdot\|)$ with $\mathbb{E}X_i = 0$ for all i . Let $\{\varepsilon_i\}$ be a sequence of independent Bernoulli r.v values independent of $\{X_i\}$. Then, for any convex increasing function Φ ,

$$\mathbb{E}\Phi\left(\frac{1}{2}\left\|\sum_{i=1}^n X_i \varepsilon_i\right\|\right) \leq \mathbb{E}\Phi\left(\left\|\sum_{i=1}^n X_i\right\|\right) \leq \mathbb{E}\Phi\left(2\left\|\sum_{i=1}^n X_i \varepsilon_i\right\|\right).$$

□

References

1. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Monographs on Statistics and Applied Probability; Chapman & Hall: London, UK, 1986; pp. x+175. [\[CrossRef\]](#)
2. Nadaraya, E.A. *Nonparametric Estimation of Probability Densities and Regression Curves*; Mathematics and Its Applications (Soviet Series); Kluwer Academic Publishers Group: Dordrecht, The Netherlands, 1989; Volume 20, pp. x+213. [\[CrossRef\]](#)
3. Härdle, W. *Applied Nonparametric Regression*; Econometric Society Monographs; Cambridge University Press: Cambridge, UK, 1990; Volume 19, pp. xvi+333. [\[CrossRef\]](#)
4. Wand, M.P.; Jones, M.C. *Kernel Smoothing*; Monographs on Statistics and Applied Probability; Chapman and Hall, Ltd.: London, UK, 1995; Volume 60, pp. xii+212. [\[CrossRef\]](#)
5. Eggermont, P.P.B.; LaRiccia, V.N. *Maximum Penalized Likelihood Estimation. Density Estimation*; Springer Series in Statistics; Springer: New York, NY, USA, 2001; Volume I, pp. xviii+510.
6. Devroye, L.; Lugosi, G. *Combinatorial Methods in Density Estimation*; Springer Series in Statistics; Springer: New York, NY, USA, 2001; pp. xii+208. [\[CrossRef\]](#)
7. Ripley, B.D. Spatial statistics: Developments 1980–1983. *Internat. Statist. Rev.* **1984**, *52*, 141–150. [\[CrossRef\]](#)
8. Rosenblatt, M. *Stationary Sequences and Random Fields*; Birkhäuser Boston, Inc.: Boston, MA, USA, 1985; p. 258. [\[CrossRef\]](#)
9. Guyon, X. *Random Fields on a Network*; Probability and Its Applications (New York); Springer: New York, NY, USA, 1995; pp. xii+255.
10. Cressie, N.A.C. *Statistics for Spatial Data*, revised ed.; Wiley Classics Library, John Wiley & Sons, Inc.: New York, NY, USA, 2015; pp. xx+900.
11. Tran, L.T. Kernel density estimation on random fields. *J. Multivar. Anal.* **1990**, *34*, 37–53. [\[CrossRef\]](#)
12. Tran, L.T.; Yakowitz, S. Nearest neighbor estimators for random fields. *J. Multivar. Anal.* **1993**, *44*, 23–46. [\[CrossRef\]](#)
13. Biau, G.; Cadre, B. Nonparametric spatial prediction. *Stat. Inference Stoch. Process.* **2004**, *7*, 327–349. [\[CrossRef\]](#)
14. Dabo-Niang, S.; Yao, A.F. Kernel spatial density estimation in infinite dimension space. *Metrika* **2013**, *76*, 19–52. [\[CrossRef\]](#)
15. Ndiaye, M.; Dabo-Niang, S.; Ngom, P. Nonparametric prediction for spatial dependent functional data under fixed sampling design. *Rev. Colomb. Estadíst.* **2022**, *45*, 391–428. [\[CrossRef\]](#)
16. Hoeffding, W. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **1948**, *19*, 293–325. [\[CrossRef\]](#)
17. Stute, W. Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **1993**, *21*, 146–156. [\[CrossRef\]](#)
18. Arcones, M.A.; Wang, Y. Some new tests for normality based on U -processes. *Statist. Probab. Lett.* **2006**, *76*, 69–82. [\[CrossRef\]](#)
19. Giné, E.; Mason, D.M. Laws of the iterated logarithm for the local U -statistic process. *J. Theoret. Probab.* **2007**, *20*, 457–485. [\[CrossRef\]](#)
20. Giné, E.; Mason, D.M. On local U -statistic processes and the estimation of densities of functions of several sample variables. *Ann. Statist.* **2007**, *35*, 1105–1145. [\[CrossRef\]](#)
21. Schick, A.; Wang, Y.; Wefelmeyer, W. Tests for normality based on density estimators of convolutions. *Statist. Probab. Lett.* **2011**, *81*, 337–343. [\[CrossRef\]](#)
22. Joly, E.; Lugosi, G. Robust estimation of U -statistics. *Stoch. Process. Appl.* **2016**, *126*, 3760–3773. [\[CrossRef\]](#)
23. Lee, S.; Linton, O.; Whang, Y.J. Testing for stochastic monotonicity. *Econometrica* **2009**, *77*, 585–602. [\[CrossRef\]](#)
24. Ghosal, S.; Sen, A.; van der Vaart, A.W. Testing monotonicity of regression. *Ann. Statist.* **2000**, *28*, 1054–1082. [\[CrossRef\]](#)
25. Abrevaya, J.; Jiang, W. A nonparametric approach to measuring and testing curvature. *J. Bus. Econom. Statist.* **2005**, *23*, 1–19. [\[CrossRef\]](#)
26. Nolan, D.; Pollard, D. U -processes: Rates of convergence. *Ann. Statist.* **1987**, *15*, 780–799. [\[CrossRef\]](#)
27. Arcones, M.A.; Giné, E. Limit theorems for U -processes. *Ann. Probab.* **1993**, *21*, 1494–1542. [\[CrossRef\]](#)

28. Sherman, R.P. Maximal inequalities for degenerate U -processes with applications to optimization estimators. *Ann. Statist.* **1994**, *22*, 439–459. [[CrossRef](#)]
29. de la Peña, V.H.; Giné, E. *Decoupling. From Dependence to Independence, Randomly Stopped Processes. U-Statistics and Processes. Martingales and Beyond*; Probability and Its Applications (New York); Springer: New York, NY, USA, 1999; pp. xvi+392. [[CrossRef](#)]
30. Halmos, P.R. The theory of unbiased estimation. *Ann. Math. Stat.* **1946**, *17*, 34–43. [[CrossRef](#)]
31. von Mises, R. On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Stat.* **1947**, *18*, 309–348. [[CrossRef](#)]
32. Yoshihara, K.i. Limiting behavior of U -statistics for stationary, absolutely regular processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **1976**, *35*, 237–252. [[CrossRef](#)]
33. Borovkova, S.; Burton, R.; Dehling, H. Limit theorems for functionals of mixing processes with applications to U -statistics and dimension estimation. *Trans. Amer. Math. Soc.* **2001**, *353*, 4261–4318. [[CrossRef](#)]
34. Denker, M.; Keller, G. On U -statistics and v. Mises' statistics for weakly dependent processes. *Z. Wahrsch. Verw. Gebiete* **1983**, *64*, 505–522. [[CrossRef](#)]
35. Leucht, A. Degenerate U - and V -statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli* **2012**, *18*, 552–585. [[CrossRef](#)]
36. Leucht, A.; Neumann, M.H. Degenerate U - and V -statistics under ergodicity: Asymptotics, bootstrap and applications in statistics. *Ann. Inst. Statist. Math.* **2013**, *65*, 349–386. [[CrossRef](#)]
37. Bouzebda, S.; Nemouchi, B. Weak-convergence of empirical conditional processes and conditional U -processes involving functional mixing data. *Stat. Inference Stoch. Process. Appear* **2022**, 1–56. [[CrossRef](#)]
38. Bouzebda, S.; Nezzal, A.; Zari, T. Uniform consistency for functional conditional U -statistics using delta-sequences. *Mathematics* **2022**, *24*, 3745.
39. Soukariéh, I.; Bouzebda, S. Exchangeably Weighted Bootstraps of General Markov U -Process. *Mathematics* **2022**, *10*, 3745. [[CrossRef](#)]
40. Bouzebda, S.; Soukariéh, I. Renewal type bootstrap for increasing degree U -process of a Markov chain. *J. Multivar. Anal.* **2022**, *195*, 105143.
41. Bouzebda, S.; Soukariéh, I. Renewal type bootstrap for U -process Markov chains. *Markov Process. Related Fields* **2022**, *13*, 1–50.
42. Frees, E.W. Infinite order U -statistics. *Scand. J. Statist.* **1989**, *16*, 29–45.
43. Rempala, G.; Gupta, A. Weak limits of U -statistics of infinite order. *Random Oper. Stochastic Equ.* **1999**, *7*, 39–52. [[CrossRef](#)]
44. Heilig, C.; Nolan, D. Limit theorems for the infinite-degree U -process. *Statist. Sinica* **2001**, *11*, 289–302.
45. Song, Y.; Chen, X.; Kato, K. Approximating high-dimensional infinite-order U -statistics: Statistical and computational guarantees. *Electron. J. Stat.* **2019**, *13*, 4794–4848. [[CrossRef](#)]
46. Peng, W.; Coleman, T.; Mentch, L. Rates of convergence for random forests via generalized U -statistics. *Electron. J. Stat.* **2022**, *16*, 232–292. [[CrossRef](#)]
47. Faivishevsky, L.; Goldberger, J. ICA based on a Smooth Estimation of the Differential Entropy. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–10 December 2008; Koller, D., Schuurmans, D., Bengio, Y., Bottou, L., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2008; Volume 21.
48. Liu, Q.; Lee, J.; Jordan, M. A Kernelized Stein Discrepancy for Goodness-of-fit Tests. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; Balcan, M.F., Weinberger, K.Q., Eds.; PMLR: New York, New York, USA, 2016; Volume 48, pp. 276–284.
49. Cléménçon, S. On U -processes and clustering performance. In Proceedings of the Advances in Neural Information Processing Systems, Granada, Spain, 12–15 December 2011; Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2011; Volume 24.
50. Borovskikh, Y.V. *U-Statistics in Banach Spaces*; VSP: Utrecht, The Netherlands, 1996; pp. xii+420.
51. Koroljuk, V.S.; Borovskikh, Y.V. *Theory of U-Statistics*; Mathematics and Its Applications; Kluwer Academic Publishers Group: Dordrecht, The Netherlands, 1994; Volume 273, pp. x+552.
52. Lee, A.J. *U-Statistics. Theory and Practice*; Statistics: Textbooks and Monographs; Marcel Dekker Inc.: New York, NY, USA, 1990; Volume 110, pp. xii+302.
53. Aneiros, G.; Cao, R.; Fraiman, R.; Genest, C.; Vieu, P. Recent advances in functional data analysis and high-dimensional statistics. *J. Multivar. Anal.* **2019**, *170*, 3–9. [[CrossRef](#)]
54. Ramsay, J.O.; Silverman, B.W. *Applied Functional Data Analysis. Methods and Case Studies*; Springer Series in Statistics; Springer: New York, NY, USA, 2002; pp. x+190. [[CrossRef](#)]
55. Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis. Theory and Practice*; Springer Series in Statistics; Springer: New York, NY, USA, 2006; pp. xx+258.
56. Araujo, A.; Giné, E. *The Central Limit Theorem for Real and Banach Valued Random Variables*; Wiley Series in Probability and Mathematical Statistics; John Wiley & Sons: New York, NY, USA, 1980; pp. xiv+233.
57. Gasser, T.; Hall, P.; Presnell, B. Nonparametric estimation of the mode of a distribution of random curves. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1998**, *60*, 681–691. [[CrossRef](#)]
58. Bosq, D. *Linear Processes in Function Spaces. Theory and Applications*; Lecture Notes in Statistics; Springer: New York, NY, USA, 2000; Volume 149, pp. xiv+283. . [[CrossRef](#)]

59. Horváth, L.; Kokoszka, P. *Inference for Functional Data with Applications*; Springer Series in Statistics; Springer: New York, NY, USA, 2012; pp. xiv+422. [[CrossRef](#)]
60. Ling, N.; Vieu, P. Nonparametric modelling for functional data: Selected survey and tracks for future. *Statistics* **2018**, *52*, 934–949. [[CrossRef](#)]
61. Ferraty, F.; Laksaci, A.; Tadj, A.; Vieu, P. Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference* **2010**, *140*, 335–352. [[CrossRef](#)]
62. Bouzebda, S.; Chaouch, M. Uniform limit theorems for a class of conditional Z-estimators when covariates are functions. *J. Multivar. Anal.* **2022**, *189*, 104872. [[CrossRef](#)]
63. Kara-Zaitri, L.; Laksaci, A.; Rachdi, M.; Vieu, P. Uniform in bandwidth consistency for various kernel estimators involving functional data. *J. Nonparametr. Stat.* **2017**, *29*, 85–107. [[CrossRef](#)]
64. Attouch, M.; Laksaci, A.; Rfaaa, F. On the local linear estimate for functional regression: Uniform in bandwidth consistency. *Comm. Statist. Theory Methods* **2019**, *48*, 1836–1853. [[CrossRef](#)]
65. Ling, N.; Meng, S.; Vieu, P. Uniform consistency rate of kNN regression estimation for functional time series data. *J. Nonparametr. Stat.* **2019**, *31*, 451–468. [[CrossRef](#)]
66. Bouzebda, S.; Chaouch, M.; Laïb, N. Limiting law results for a class of conditional mode estimates for functional stationary ergodic data. *Math. Methods Statist.* **2016**, *25*, 168–195. [[CrossRef](#)]
67. Mohammedi, M.; Bouzebda, S.; Laksaci, A. The consistency and asymptotic normality of the kernel type expectile regression estimator for functional data. *J. Multivar. Anal.* **2021**, *181*, 104673. [[CrossRef](#)]
68. Bouzebda, S.; Mohammedi, M.; Laksaci, A. The k-Nearest Neighbors method in single index regression model for functional quasi-associated time series data. *Rev. Mat. Complut.* **2022**, 1–30. [[CrossRef](#)]
69. Bouzebda, S.; Nezzal, A. 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.* **2022**, *5*, 431–533. [[CrossRef](#)]
70. Didi, S.; Al Harby, A.; Bouzebda, S. Wavelet Density and Regression Estimators for Functional Stationary and Ergodic Data: Discrete Time. *Mathematics* **2022**, *10*, 3433. [[CrossRef](#)]
71. Almanjahie, I.M.; Bouzebda, S.; Kaid, Z.; Laksaci, A. Nonparametric estimation of expectile regression in functional dependent data. *J. Nonparametr. Stat.* **2022**, *34*, 250–281. [[CrossRef](#)]
72. Almanjahie, I.M.; Bouzebda, S.; Chikr Elmezouar, Z.; Laksaci, A. The functional kNN estimator of the conditional expectile: uniform consistency in number of neighbors. *Stat. Risk Model.* **2022**, *38*, 47–63. [[CrossRef](#)]
73. Stute, W. Conditional U-statistics. *Ann. Probab.* **1991**, *19*, 812–825. [[CrossRef](#)]
74. Sen, A. Uniform strong consistency rates for conditional U-statistics. *Sankhyā Ser. A* **1994**, *56*, 179–194.
75. Prakasa Rao, B.L.S.; Sen, A. Limit distributions of conditional U-statistics. *J. Theoret. Probab.* **1995**, *8*, 261–301. [[CrossRef](#)]
76. Harel, M.; Puri, M.L. Conditional U-statistics for dependent random variables. *J. Multivar. Anal.* **1996**, *57*, 84–100. [[CrossRef](#)]
77. Stute, W. Symmetrized NN-conditional U-statistics. In *Research developments in Probability and Statistics*; VSP: Utrecht, The Netherlands, 1996; pp. 231–237.
78. Fu, K.A. An application of U-statistics to nonparametric functional data analysis. *Comm. Statist. Theory Methods* **2012**, *41*, 1532–1542. [[CrossRef](#)]
79. Bouzebda, S.; Nemouchi, B. Uniform consistency and uniform in bandwidth consistency for nonparametric regression estimates and conditional U-statistics involving functional data. *J. Nonparametr. Stat.* **2020**, *32*, 452–509. [[CrossRef](#)]
80. Bouzebda, S.; Elhattab, I.; Nemouchi, B. On the uniform-in-bandwidth consistency of the general conditional U-statistics based on the copula representation. *J. Nonparametr. Stat.* **2021**, *33*, 321–358. [[CrossRef](#)]
81. Jadhav, S.; Ma, S. Kendall’s Tau for Functional Data Analysis. *arXiv* **2019**, arXiv:1912.03725.
82. Arcones, M.A.; Yu, B. Central limit theorems for empirical and U-processes of stationary mixing sequences. *J. Theoret. Probab.* **1994**, *7*, 47–71. [[CrossRef](#)]
83. Bouzebda, S.; Nemouchi, B. Central limit theorems for conditional empirical and conditional U-processes of stationary mixing sequences. *Math. Methods Statist.* **2019**, *28*, 169–207. [[CrossRef](#)]
84. Masry, E. Nonparametric regression estimation for dependent functional data: Asymptotic normality. *Stoch. Process. Appl.* **2005**, *115*, 155–177. [[CrossRef](#)]
85. Kurisu, D. Nonparametric regression for locally stationary random fields under stochastic sampling design. *Bernoulli* **2022**, *28*, 1250–1275. [[CrossRef](#)]
86. Kurisu, D. Nonparametric regression for locally stationary functional time series. *Electron. J. Stat.* **2022**, *16*, 3973–3995. [[CrossRef](#)]
87. Kurisu, D.; Kato, K.; Shao, X. Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data. *arXiv* **2021**, arXiv:2103.10720.
88. Dahlhaus, R. Fitting time series models to nonstationary processes. *Ann. Statist.* **1997**, *25*, 1–37. [[CrossRef](#)]
89. Dahlhaus, R.; Subba Rao, S. Statistical inference for time-varying ARCH processes. *Ann. Statist.* **2006**, *34*, 1075–1114. [[CrossRef](#)]
90. van Delft, A.; Eichler, M. Locally stationary functional time series. *Electron. J. Stat.* **2018**, *12*, 107–170. [[CrossRef](#)]
91. Hall, P.; Patil, P. Properties of nonparametric estimators of autocovariance for stationary random fields. *Probab. Theory Related Fields* **1994**, *99*, 399–424. [[CrossRef](#)]
92. Matsuda, Y.; Yajima, Y. Fourier analysis of irregularly spaced data on \mathbb{R}^d . *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2009**, *71*, 191–217. [[CrossRef](#)]

93. Lahiri, S.N. Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhyā* **2003**, *65*, 356–388.
94. Lahiri, S.N. *Resampling Methods for Dependent Data*; Springer Series in Statistics; Springer: New York, NY, USA, 2003; pp. xiv+374. [[CrossRef](#)]
95. Volkonskiĭ, V.A.; Rozanov, Y.A. Some limit theorems for random functions. I. *Theor. Probability Appl.* **1959**, *4*, 178–197. [[CrossRef](#)]
96. Rosenblatt, M. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **1956**, *27*, 832–837. [[CrossRef](#)]
97. Ibragimov, I.A.; Solev, V.N. A condition for the regularity of a Gaussian stationary process. *Dokl. Akad. Nauk SSSR* **1969**, *185*, 509–512.
98. Bradley, R.C. A caution on mixing conditions for random fields. *Statist. Probab. Lett.* **1989**, *8*, 489–491. [[CrossRef](#)]
99. Bradley, R.C. Some examples of mixing random fields. *Rocky Mountain J. Math.* **1993**, *23*, 495–519. [[CrossRef](#)]
100. Doukhan, P. *Mixing. Properties and examples*; Lecture Notes in Statistics; Springer: New York, NY, USA, 1994; Volume 85, pp. xii+142. [[CrossRef](#)]
101. Dedecker, J.; Doukhan, P.; Lang, G.; León, R.; Louhichi, S.; Prieur, C. *Weak Dependence: With Examples and Applications*; Lecture Notes in Statistics; Springer: New York, NY, USA, 2007; Volume 190, pp. xiv+318.
102. Lahiri, S.N.; Zhu, J. Resampling methods for spatial regression models under a class of stochastic designs. *Ann. Statist.* **2006**, *34*, 1774–1813. [[CrossRef](#)]
103. Bandyopadhyay, S.; Lahiri, S.N.; Nordman, D.J. A frequency domain empirical likelihood method for irregularly spaced spatial data. *Ann. Statist.* **2015**, *43*, 519–545. [[CrossRef](#)]
104. Kolmogorov, A.N.; Tihomirov, V.M. ε -entropy and ε -capacity of sets in functional space. *Amer. Math. Soc. Transl.* **1961**, *17*, 277–364.
105. Dudley, R.M. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* **1967**, *1*, 290–330. [[CrossRef](#)]
106. Dudley, R.M. *Uniform Central Limit Theorems*; Cambridge Studies in Advanced Mathematics; Cambridge University Press: Cambridge, UK, 1999; Volume 63, pp. xiv+436. [[CrossRef](#)]
107. van der Vaart, A.W.; Wellner, J.A. *Weak Convergence and Empirical Processes. With Applications to Statistics*; Springer Series in Statistics; Springer: New York, NY, USA, 1996; pp. xvi+508. [[CrossRef](#)]
108. Kosorok, M.R. *Introduction to Empirical Processes and Semiparametric Inference*; Springer Series in Statistics; Springer: New York, NY, USA, 2008; pp. xiv+483. [[CrossRef](#)]
109. Deheuvels, P. One bootstrap suffices to generate sharp uniform bounds in functional estimation. *Kybernetika* **2011**, *47*, 855–865.
110. Vogt, M. Nonparametric regression for locally stationary time series. *Ann. Statist.* **2012**, *40*, 2601–2633. [[CrossRef](#)]
111. Mayer-Wolf, E.; Zeitouni, O. The probability of small Gaussian ellipsoids and associated conditional moments. *Ann. Probab.* **1993**, *21*, 14–24. [[CrossRef](#)]
112. Bogachev, V.I. *Gaussian Measures*; Mathematical Surveys and Monographs; American Mathematical Society: Providence, RI, USA, 1998; Volume 62, pp. xii+433. [[CrossRef](#)]
113. Li, W.V.; Shao, Q.M. Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes: Theory and Methods*; Handbook of Statist: Amsterdam, The Netherlands, 2001; Volume 19, pp. 533–597. [[CrossRef](#)]
114. Ferraty, F.; Mas, A.; Vieu, P. Nonparametric regression on functional data: Inference and practical aspects. *Aust. N. Z. J. Stat.* **2007**, *49*, 267–286. [[CrossRef](#)]
115. Lahiri, S.N.; Kaiser, M.S.; Cressie, N.; Hsu, N.J. Prediction of spatial cumulative distribution functions using subsampling. *J. Amer. Statist. Assoc.* **1999**, *94*, 86–110. [[CrossRef](#)]
116. van der Vaart, A.W. *Asymptotic Statistics*; Cambridge Series in Statistical and Probabilistic Mathematics; Cambridge University Press: Cambridge, UK, 1998; Volume 3, pp. xvi+443. [[CrossRef](#)]
117. Mason, D.M. Proving consistency of non-standard kernel estimators. *Stat. Inference Stoch. Process.* **2012**, *15*, 151–176. [[CrossRef](#)]
118. Bellet, A.; Habrard, A.; Sebban, M. A Survey on Metric Learning for Feature Vectors and Structured Data. *arXiv* **2013**, arXiv:1306.6709.
119. Cléménçon, S.; Colin, I.; Bellet, A. Scaling-up empirical risk minimization: Optimization of incomplete U -statistics. *J. Mach. Learn. Res.* **2016**, *17*, 76.
120. Jin, R.; Wang, S.; Zhou, Y. Regularized Distance Metric Learning: Theory and Algorithm. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009; Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C., Culotta, A., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2009; Volume 22.
121. Bellet, A.; Habrard, A. Robustness and generalization for metric learning. *Neurocomputing* **2015**, *151*, 259–267. [[CrossRef](#)]
122. Cao, Q.; Guo, Z.C.; Ying, Y. Generalization bounds for metric and similarity learning. *Mach. Learn.* **2016**, *102*, 115–132. [[CrossRef](#)]
123. Cléménçon, S.; Robbiano, S. The TreeRank Tournament algorithm for multipartite ranking. *J. Nonparametr. Stat.* **2015**, *27*, 107–126. [[CrossRef](#)]
124. Cléménçon, S.; Robbiano, S.; Vayatis, N. Ranking data with ordinal labels: Optimality and pairwise aggregation. *Mach. Learn.* **2013**, *91*, 67–104. [[CrossRef](#)]
125. Dudley, R.M. A course on empirical processes. In *École d'été de Probabilités de Saint-Flour, XII—1982*; Lecture Notes in Math; Springer: Berlin, Germany, 1984; Volume 1097, pp. 1–142. [[CrossRef](#)]
126. Polonik, W.; Yao, Q. Set-indexed conditional empirical and quantile processes based on dependent data. *J. Multivar. Anal.* **2002**, *80*, 234–255. [[CrossRef](#)]

127. Stute, W. Universally consistent conditional U -statistics. *Ann. Statist.* **1994**, *22*, 460–473. [[CrossRef](#)]
128. Stute, W. L^p -convergence of conditional U -statistics. *J. Multivar. Anal.* **1994**, *51*, 71–82. [[CrossRef](#)]
129. Maillot, B.; Viallon, V. Uniform limit laws of the logarithm for nonparametric estimators of the regression function in presence of censored data. *Math. Methods Statist.* **2009**, *18*, 159–184. [[CrossRef](#)]
130. Kohler, M.; Máthé, K.; Pintér, M. Prediction from randomly right censored data. *J. Multivar. Anal.* **2002**, *80*, 73–100. [[CrossRef](#)]
131. Carbonez, A.; Györfi, L.; van der Meulen, E.C. Partitioning-estimates of a regression function under random censoring. *Statist. Decis.* **1995**, *13*, 21–37. [[CrossRef](#)]
132. Brunel, E.; Comte, F. Adaptive nonparametric regression estimation in presence of right censoring. *Math. Methods Statist.* **2006**, *15*, 233–255.
133. Kaplan, E.L.; Meier, P. Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.* **1958**, *53*, 457–481. [[CrossRef](#)]
134. Bouzebda, S.; El-hadjali, T. Uniform convergence rate of the kernel regression estimator adaptive to intrinsic dimension in presence of censored data. *J. Nonparametr. Stat.* **2020**, *32*, 864–914. [[CrossRef](#)]
135. Datta, S.; Bandyopadhyay, D.; Satten, G.A. Inverse probability of censoring weighted U -statistics for right-censored data with an application to testing hypotheses. *Scand. J. Stat.* **2010**, *37*, 680–700. [[CrossRef](#)]
136. Stute, W.; Wang, J.L. Multi-sample U -statistics for censored data. *Scand. J. Statist.* **1993**, *20*, 369–374.
137. Chen, Y.; Datta, S. Adjustments of multi-sample U -statistics to right censored data and confounding covariates. *Comput. Statist. Data Anal.* **2019**, *135*, 1–14. [[CrossRef](#)]
138. Yuan, A.; Giurcanu, M.; Luta, G.; Tan, M.T. U -statistics with conditional kernels for incomplete data models. *Ann. Inst. Statist. Math.* **2017**, *69*, 271–302. [[CrossRef](#)]
139. Földes, A.; Rejtő, L. A LIL type result for the product limit estimator. *Z. Wahrsch. Verw. Gebiete* **1981**, *56*, 75–86. [[CrossRef](#)]
140. Bouzebda, S.; El-hadjali, T.; Ferfache, A.A. Uniform in bandwidth consistency of conditional U -statistics adaptive to intrinsic dimension in presence of censored data. *Sankhya A* **2022**, 1–59. [[CrossRef](#)]
141. Hall, P. Asymptotic properties of integrated square error and cross-validation for kernel estimation of a regression function. *Z. Wahrsch. Verw. Gebiete* **1984**, *67*, 175–196. [[CrossRef](#)]
142. Härdle, W.; Marron, J.S. Optimal bandwidth selection in nonparametric regression function estimation. *Ann. Statist.* **1985**, *13*, 1465–1481. [[CrossRef](#)]
143. Rachdi, M.; Vieu, P. Nonparametric regression for functional data: Automatic smoothing parameter selection. *J. Statist. Plann. Inference* **2007**, *137*, 2784–2801. [[CrossRef](#)]
144. Benhenni, K.; Ferraty, F.; Rachdi, M.; Vieu, P. Local smoothing regression with functional data. *Comput. Statist.* **2007**, *22*, 353–369. [[CrossRef](#)]
145. Shang, H.L. Bayesian bandwidth estimation for a functional nonparametric regression model with mixed types of regressors and unknown error density. *J. Nonparametr. Stat.* **2014**, *26*, 599–615. [[CrossRef](#)]
146. Li, Q.; Maasoumi, E.; Racine, J.S. A nonparametric test for equality of distributions with mixed categorical and continuous data. *J. Econom.* **2009**, *148*, 186–200. [[CrossRef](#)]
147. Horowitz, J.L.; Spokoiny, V.G. An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica* **2001**, *69*, 599–631. [[CrossRef](#)]
148. Gao, J.; Gijbels, I. Bandwidth selection in nonparametric kernel testing. *J. Am. Statist. Assoc.* **2008**, *103*, 1584–1594. [[CrossRef](#)]
149. Yu, B. Rates of convergence for empirical processes of stationary mixing sequences. *Ann. Probab.* **1994**, *22*, 94–116. [[CrossRef](#)]
150. Bernstein, S. Sur l’extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann.* **1927**, *97*, 1–59. [[CrossRef](#)]
151. Giné, E.; Zinn, J. Some limit theorems for empirical processes. *Ann. Probab.* **1984**, *12*, 929–998. [[CrossRef](#)]
152. Bouzebda, S.; Soukariéh, I. Weak Convergence of the Conditional U -statistics for Locally Stationary Functional Time Series. *Stat. Inference Stoch. Process.* **2022**.
153. Masry, E. Multivariate local polynomial regression for time series: Uniform strong consistency and rates. *J. Time Ser. Anal.* **1996**, *17*, 571–599. [[CrossRef](#)]
154. Rio, E. Inequalities and Limit Theorems for Weakly Dependent Sequences. Available online: <https://cel.hal.science/cel-00867106/> (accessed on 20 October 2022).
155. de la Peña, V.H. Decoupling and Khintchine’s inequalities for U -statistics. *Ann. Probab.* **1992**, *20*, 1877–1892. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.