Article

Exchangeably Weighted Bootstraps of General Markov *U*-Process

Inass Soukarieh [†] and Salim Bouzebda *,[†]

Laboratory of Applied Mathematics of Compiègne (LMAC), Université de Technologie de Compiègne, 60200 Compiègne, France
* Correspondence: salim.bouzebda@utc.fr
† These authors contributed equally to this work.

Abstract: We explore an exchangeably weighted bootstrap of the general function-indexed empirical *U*-processes in the Markov setting, which is a natural higher-order generalization of the weighted bootstrap empirical processes. As a result of our findings, a considerable variety of bootstrap resampling strategies arise. This paper aims to provide theoretical justifications for the exchangeably weighted bootstrap consistency in the Markov setup. General structural conditions on the classes of functions (possibly unbounded) and the underlying distributions are required to establish our results. This paper provides the first general theoretical study of the bootstrap of the empirical *U*-processes in the Markov setting. Potential applications include the symmetry test, Kendall's tau and the test of independence.

1. Introduction

U-statistics are a class of estimators, initially explored in association with unbiased estimators by [1] and officially introduced by [2], and are defined as follows: let $\{X_i\}_{i=1}^{\infty}$ be a sequence of random variables defined on a measurable space $(E, \mathscr{E})$, and let $h : E^m \to \mathbb{R}$ be a measurable function, the *U*-statistics of order m and kernel h based on the sequence $\{X_i\}$ are

$$U_n(h) = \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m) \in I_n^m} h(X_{i_1}, \ldots, X_{i_m}), \quad n \geq m,$$

where

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

The empirical variance, Gini's mean difference or Kendall's rank correlation coefficient are common examples of *U*-estimators, while a classical test based on a *U*-statistic is Wilcoxon's signed rank test for the hypothesis of the location at zero (see, e.g., [3], Example 12.4). The authors in [1,2,4] provided, amongst others, the first asymptotic results for the case in which the underlying random variables have independent and identical distributions. Extensive literature works have treated the theory of *U*-statistics, for instance, see [5–8], etc. Complex statistical issues are also amenable to being solved using *U*-processes. Examples include tests for goodness-of-fit, nonparametric regression and density estimation. *U*-processes are a set of *U*-statistics that are indexed by a family of kernels. *U*-processes might be viewed as infinite-dimensional variants of *U*-statistics with a single kernel function or as nonlinear stochastic extensions of empirical processes. Both thoughts have the following advantages: first, considering a large group of statistics rather than a single statistic is

more statistically interesting. Second, we may use ideas from the theory of empirical processes to construct limit or approximation theorems for U-processes. Nevertheless, achieving results in U-processes is not easy. Extending U-statistics to U-processes necessitates a significant effort and distinct methodologies; generalizing empirical processes to U-processes is quite challenging, especially when U-processes are presented in the stationary setting. We highlight that the U-processes are used often in statistics, such as when higher order terms are a part of von Mises expansions. Particularly, the study of estimators (including function estimators) with various smoothness degrees involves U-statistics. For instance, Ref. [9] applied almost-sure uniform bounds for $\mathbb{P}$-canonical U-processes to analyze the product limit estimator for truncated data. Two new tests for normality based on U-processes were also presented in [10]. Inspired by [11–13], they developed other tests for normality that employed weighted L_1-distances between the standard normal density and local U-statistics based on standardized observations as test statistics. Estimating the mean of multivariate functions in the case of possibly heavy-tailed distributions was explored by [14]; they presented the median-of-means too, and both explorations were based on U-statistics. Moreover, other researchers emphasized the importance of U-processes; refs. [15–17] used them for testing qualitative features of functions in nonparametric statistics, ref. [18] represented the cross-validation for density estimation using U-statistics, in addition to [6,7,19], where the authors established limiting distributions of M-estimators. Since then, this discipline has made significant advancements, and the results have been broadly interpreted. Asymptotic behaviors were demonstrated under weak dependence assumptions, for example, in the works of [20–22] or more recently in [23] as well as more generally in [24,25]. However, in practice, explicit computation is not always possible due to the complexity of the U-processes' limiting distributions or their functionals. We suggest a general bootstrap of the U-processes in the Markov setting to solve this issue, which is a challenging problem. The concept of the bootstrap, given by [26], in the case of independent and identically distributed (iid) random variables, is to resample from an original sample of observations of an unknown marginal distribution function $F(x)$, $X_1, \ldots, X_n$, a new i.i.d sample $X_1^*, \ldots, X_n^*$ with the marginal distribution function $F_n(x)$, which represents the empirical distribution function constructed from the original sample. Moreover, it is commonly known that the bootstrap approach gives a better approximation to the statistic's distribution, mainly when the sample size is small [27]. Bootstraps for U-statistics of independent observations were studied by [28–31]. However, the bootstrap technique is not the same for dependent variables because the dependence structure cannot be conserved in the new sample. For this reason, other *blockwise bootstrap methods* were introduced, aiming to keep the structure of dependence. Among those methods, we can cite the *circular block bootstrap* introduced by [32] and the *nonoverlapping block bootstrap* introduced by [33]. In [34], the authors proposed a bootstrap method related to the weakly dependent stationary observation, the *stationary bootstrap*. This latter can be seen as an expansion of the circular block bootstrap, where a random variable, such as a geometric random variable, can be used for the block length. It is important to note that Efron's initial bootstrap formulation (see [26]) had a few flaws. To be more precise, certain observations might be sampled several times while others might not be at all. A more generalized version of the bootstrap, the *weighted* bootstrap, was developed to get around this issue and was also demonstrated to be computationally more appealing in some applications. This resampling strategy was initially described in [35] and thoroughly investigated by [28], who coined the name "weighted bootstrap". For example, Bayesian bootstrap when the weighted vector

$$(\xi_{n1}, \ldots, \xi_{nn}) = (M_{n1}, \ldots, M_{nn}),$$

is equal to the vector of n spacings of $n-1$ ordered uniform $(0,1)$ random variables in distributions, that is, $(M_{n1}, \ldots, M_{nn})$ follows a Dirichlet distribution of parameters $(n; 1, \ldots, 1)$. For more details, see [36]. This diversity of resampling approaches necessitates the use of a uniform approach, commonly known as general weighted resampling, which was first

described by [37] and has since been developed by [38,39]. In [40], the authors investigated the almost-sure rate of convergence of strong approximation for the weighted bootstrap process by a sequence of Brownian bridge processes; refer to [41] for the multivariate setting and [42] for recent references. The concept of the generalized bootstrap, introduced by [37], was extended to the class of nondegenerate U-statistics of degree two and the corresponding Studentized U-statistics by [43]; refer to [44,45]. In [46], the author generalized this theory for a higher order. In his work, he developed a multiplier inequality of a U-process for i.i.d. random variables. We mention that the multiplier processes' theory is directly and strongly related to the symmetrization inequalities investigated by [6,7].

This paper aims to investigate the exchangeable bootstrap for U-processes in the same way that [46] did but without the restriction of the independence setting. The previous reference focused on U-processes in an independent framework, whereas this paper considers U-processes in the dependent setting of Markov chains. We believe we are the first to present a successful consideration in this general context. We combine the techniques of the renewal bootstrap with the randomly weighted bootstrap in a nontrivial way. We mention a connection between moving-blocks bootstrap and its modification, matched-block bootstrap, at this point. Instead of artificially splitting a sample into fixed-size blocks and then resampling them, the latter seeks to match the blocks to create a smoother transition; for more information, see [47]. The main difficulties in proving Theorem 3 are due to the random size of the resampled blocks. This randomness generates a problem with the random stopping times, which cannot be removed by replacing a random stopping time with its expectation. In the present setting, the bootstrap random variables are generated by resampling from a random number of blocks. One can think that using the conditioning arguments can overcome the problem, but the answer is negative. Our proof uses some arguments from [46,47] by verifying bootstrap stochastic equicontinuity by comparing it to the original process in a similar way as in [48]. However, as we shall see later, integrating concepts from these papers is not enough to solve the problem. To deal with U-processes in the Markov framework, sophisticated mathematical derivations are necessary. We present the first complete theoretical justification of the bootstrap consistency. This justification requires the efficient use of large sample theoretical approaches established for U-empirical processes.

The rest of this paper is organized as follows. Section 2 is devoted to the introduction of the Markov framework, the U-process, the bootstrap weights and the definitions needed in our work. In Section 3, we recall the necessary ingredient for U-statistics and U-processes in the Markov setting. Furthermore, we provide some asymptotic results including the weak convergence of U-processes in Theorem 1. In Section 4, we derive the main results concerning the bootstrap of the U-processes. In Section 5, we collect some examples of weighted U-statistics. Some concluding remarks and possible future developments are relegated to Section 6. To prevent the interruption of the flow of the presentation, all proofs are gathered in Section 7. Appendix A contains a few pertinent technical findings and proofs.

2. Notation and Definitions

In what follows, we aim to properly define our settings. For this reason, we have collected the definitions and notation needed.

2.1. Markov Chain

Let $X = (X_n)_{n \in \mathbb{N}}$ be an homogeneous ψ-*irreducible* Markov chain, that means that the chain has stationary transition probabilities, defined on a measurable space $(E, \mathscr{E})$, where $\mathscr{E}$ is a separable σ-algebra. Let $\pi(x, dy)$ be the transition probability and $\nu = \nu(i)_{i>0}$ the initial probability. Therefore, we denote by $\mathbb{P}_\nu$ or just $\mathbb{P}$ the probability measure for $\mathbb{P} = (\pi, \nu)$. Likewise, $\mathbb{E}_\nu$ denotes the integration with respect to $\mathbb{P}_\nu$. In our framework, let $\mathbb{P}_x$ be a probability measure such that $X_0 = x$, $X_0 \in E$ and $\mathbb{E}_x(\cdot)$ is the $\mathbb{P}_x$-expectation. We further assume that the Markov chain is Harris positive recurrent with an atom A.

Definition 1 (Harris recurrent). *A Markov chain $X = (X_n)_{n \in \mathbb{N}}$ is said to be Harris recurrent if there exists a σ-finite measure such that, for ψ a positive measure on a countable generated measurable space $(E, \mathscr{E})$, $\psi(E) > 0$ and if for all $B \in E$ with $\psi(B) > 0$*

$$P_x(\cup_{i=1}^{\infty}(X_i \in B)) = 1 \ \text{for any } x \in E.$$

Recall that a chain is positive Harris recurrent and aperiodic if and only if it is ergodic ([49] Proposition 6.3), i.e., there exists a probability measure π, called the stationary distribution, such that, in total variation distance,

$$\lim_{n \to +\infty} \|\mathbb{P}^n(x, \cdot) - \pi\|_{\text{tv}} = 0.$$

Definition 2 (Small sets). *A set $S \in \mathscr{E}$ is said to be Ψ-small if there exists $\delta > 0$, a positive probability measure Ψ supported by S and an integer $m \in \mathbb{N}^*$, such that*

$$\forall x \in S, \ B \in \mathscr{E}, \ \mathbb{P}^m(x, B) \geq \delta \ \Psi(B). \tag{1}$$

Definition 3. *Let $(X_n)_{n \geq 1}$ be a Markov chain taking value in $(E, \mathscr{E})$. We say that $(X_n)_{n \geq 1}$ is positive recurrent if*

1. *$(X_n)_{n \geq 1}$ is (A, p, ν, m) recurrent (or Harris recurrent if $\mathscr{E}$ is countably generated), where $A \in \mathscr{E}$ is a set, $0 < p < 1$, m is an integer and ν is a probability measure.*
2. *$\sup\limits_{x \in A} \mathbb{E}_x(\mathscr{T}_0) < \infty$, where $\mathscr{T}_0$ is the hitting time of A by the m-step chain, roughly speaking, $\mathscr{T}_0 = \min\{i \geq 1 : X_{i,m} \in A\}$.*

Definition 4. *A ψ-irreducible aperiodic chain X is called regenerative or atomic if there exists a measurable set $\mathbf{A}$ called an atom, in such a way that $\psi(\mathbf{A}) > 0$ and for all $(x, y) \in \mathbf{A}^2$ we have $\mathbb{P}(x, \cdot) = \mathbb{P}(y, \cdot)$. Roughly speaking, an atom is a set on which the transition probabilities are the same. If a finite number of states or subsets are visited from the chain, then any state or any subset of the states is actually an atom.*

Definition 5 (Aperiodicity). *Assuming ψ-irreducibility, there exists $d' \in \mathbb{N}^*$ and disjoints sets $D_1, \ldots, D_{d'}$ (set $D_{d'+1} = D_1$) positively weighted by ψ such that*

$$\psi(E \setminus \cup_{1 \leqslant i \leqslant d'} D_i) = 0$$

and

$$\forall x \in D_i, \mathbb{P}(x, D_{i+1}) = 1.$$

The period of the chain is the greatest common divisor d of such integers, it is said to be aperiodic if $d = 1$.

Definition 6 (Irreducibility). *The chain is ψ-irreducible if there exists a σ-finite measure ψ such that, for all set $B \in \mathscr{E}$, when $\psi(B) > 0$, for any $x \in E$, there exists $n > 0$ such that $\mathbb{P}^n(x, B) > 0$.*

One of the most important properties of Harris recurrent Markov chains is the existence of an invariant distribution which we is called μ (a limiting probability distribution, also called occupation measure). Furthermore, Harris recurrent Markov chains can always be embedded in a certain Markov chain on an extended sample space with a recurrent atom. The existence of a recurrent atom A gives an immediate consequence for the construction of a regenerative extension of this chain. The time that the chain hits a given atom (recurrent state) is seen as the regenerative time. In [50,51], the authors give the construction of such a regenerative extension. The development of a regenerative extension makes the use of regenerative techniques possible in order to study this type of Markov chain. As we mentioned above, we assume in this work that the Harris recurrent chain is atomic, i.e., the set which is infinitely almost sure is well-defined and accessible, this set A is called

an atom. By definition, an atom A is a set in $\mathscr{E}$, where $\mu(A) > 0$, and for all $x, y \in A$, $\pi(x, \cdot) = \pi(y, \cdot)$. Let $\mathbb{P}_{\mathbf{A}}$ (respectively, $\mathbb{E}_{\mathbf{A}}$) be the probability measure on the underlying space such that $x \in \mathbf{A}$ (respectively, the $\mathbb{P}_{\mathbf{A}}$-expectation).

The conditions imposed on the Markov chain ensure that the defined atom A (or the constructed one in the case of a nonatomic chain) is one recurrent class, and let us define the following terms.

2.1.1. Hitting Times

Define $\mathscr{T}_j : E \to \mathbb{N} \cup \{\infty\}$ by

$$
\begin{aligned}
\mathscr{T}_0 &:= \inf\{n \geq 0 : X_n \in A\}, \\
\mathscr{T}_j &:= \inf\{n \geq \mathscr{T}_{j-1} : X_n \in A\}.
\end{aligned}
\tag{2}
$$

A well-known property of the hitting time is that for all $j \in \mathbb{N}$, $\mathscr{T}_j < \infty$, $\mathbb{P}_\nu - a.s$ ([52], chap. I14).

2.1.2. Renewal Times

Using the hitting times, we can define the renewal times as

$$
\begin{aligned}
\tau_0 &:= \mathscr{T}_0 + 1, \\
\tau(j) &:= \mathscr{T}_j - \mathscr{T}_{j-1}.
\end{aligned}
\tag{3}
$$

Similar to the regenerative process, the sequence of renewal times $\{\tau(j)\}_{j=1}^{\infty}$ is i.i.d. and it is independent of the choice of the initial probability. All over this work, we set $\tau = \tau(1)$ and $\alpha = \mathbb{E}_A(\tau)$.

Definition 7 (Strong Markov property). *Let $(X_n)_{n \geq 0}$ be a Markov chain, with T the stopping time of $(X_n)_{n \geq 0}$. Then, conditionally on $T < \infty$ and $X_T = i$, $(X_{T+n})_{n \geq 0}$ is a sequence of a Markov chain and is independent of $X_0, \ldots, X_T$.*

2.1.3. Regenerative Blocks

Let $l_n := \max\{j : \sum_{i=0}^{j} \tau(j) \leq n\}$ be the number of visits to the atom A. Using the strong property of a Markov chain, it is possible to divide the given sample $(X_1, \ldots, X_n)$ into a sequence of blocks $\{\mathbf{B}_j\}_{j=0}^{l_n}$ such that:

$$
\begin{aligned}
\mathbf{B}_0 &= \left\{ X_1, \cdots, X_{\mathscr{T}_0} \right\}, \\
\mathbf{B}_j &= \left\{ X_{\mathscr{T}_{j-1}+1}, \ldots, X_{\mathscr{T}_j} \right\} \text{ in } \mathbb{T} = \bigcup_{n=1}^{\infty} E^n, \text{for all } j = 1, \cdots, l_n, \\
\mathbf{B}_{l_n}^{(n)} &= \left\{ X_{\mathscr{T}_{l_n-1}+1}, \ldots, X_n \right\},
\end{aligned}
\tag{4}
$$

where l_n is the total number of blocks. The length of each block is denoted by

$$
l(\mathbf{B}_j) := \mathscr{T}_j - \mathscr{T}_{j-1}.
$$

2.2. Exchangeable Weights

In what follows, ξ represents a real-valued random variable, ξ_i are independent from (X_i). For $1 \leq p < \infty$, we denote the p-norm by

$$
\|\xi\|_p = (\mathbb{E}(|\xi|^p))^{1/p}.
$$

Definition 8 (Exchangeability). *Let $\xi_{n1}, \ldots, \xi_{nn}$ be a sequence of random variables with joint distribution $\mathbb{P}_\xi$ and let $\Sigma(n)$ be the group of all permutations acting on $\{1, \ldots, n\}$. We say that $\xi_{n1}, \ldots, \xi_{nn}$ is exchangeable if, for all $\sigma(i) \in \Sigma(n)$,*

$$\mathbb{P}_{\bar{\xi}}(\xi_{n1}, \ldots, \xi_{nn}) = \mathbb{P}_{\bar{\xi}}(\xi_{n\sigma(1)}, \ldots, \xi_{n\sigma(n)}).$$

Assuming the following:

(A1) $(\xi_1, \ldots, \xi_n)$ are exchangeable non-negative, symmetric and for all n

$$\sum_{i=1}^{n} \xi_i = n.$$

(A2) $\frac{1}{n} \max_{1 \leq i \leq n} (\xi_i - 1)^2 \to 0$ in $\mathbb{P}_{\bar{\xi}}$-probability which is satisfied by the assumption of the moment

$$\sup_n \|\xi_1\|_{2m,1} < \infty.$$

(A3) There exists $c > 0$ such that, in $\mathbb{P}_{\bar{\xi}}$-probability,

$$\frac{1}{n} \sum_{i=1}^{n} (\xi_i - 1)^2 \to c^2 > 0.$$

(A4) Assume

$$\lim_{\lambda \to \infty} \overline{\lim_{t \geq \lambda}} \, t^2 \mathbb{P}_{\bar{\xi}}(\xi_1 \geq t) = 0.$$

2.3. The U-Process Framework

Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of random variables with values in a measurable space $(E, \mathscr{E})$. Let $h : E^m \to \mathbb{R}$ be a measurable function symmetric in its arguments. The *U*-statistic of order (or degree) m and kernel $h(\cdot)$ is defined as:

$$U_n(h) = \binom{n}{m}^{-1} \sum_{(i_1, \ldots, i_m) \in I_n^m} h(X_{i_1}, \ldots, X_{i_m}), \text{ for } n \geq m. \tag{5}$$

Accordingly, a *U*-process is the collection $\{U_n(h) : h \in \mathscr{F}\}$, where $\mathscr{F}$ is the class of kernels $h(\cdot)$ of m variables. The decoupling inequality of *U*-statistics and *U*-processes plays a central role in the latest developments in the asymptotic theory. As a result, the decoupling inequality can give a relation between the quantities

$$\mathbb{E}\Phi\left(\left|\sum_{I_n^m} h(X_{i_1}, \ldots, X_{i_m})\right|\right) \text{ and } \mathbb{E}\Phi\left(\left|\sum_{I_n^m} h\left(X_{i_1}^1, \ldots, X_{i_m}^m\right)\right|\right),$$

where $\Phi(\cdot)$ is a non-negative function and $\{X_i^k\}, k = 1, \ldots, m$ are independent copies of the original sequence $\{X_i\}$. One of the useful reasons for decoupling is randomization, which is frequently used in the study of the asymptotic theory of *U*-statistics, and was studied by [6,7]. The main idea of randomization is to compare the tail probabilities or moments of the original *U*-statistic or process, $\sum_{I_n^m} h(X_{i_1}, \ldots, X_{i_m})$, with the tail probabilities or moments of the statistic

$$\sum_{I_n^m} \varepsilon_{i_1} \ldots \varepsilon_{i_r} h(X_{i_1}, \ldots, X_{i_m}),$$

where ε_i are independent Rademacher variables, independent from X_i, $1 \leq r \leq m$ and the variables depend on the degree of degeneracy (centering) of the kernel $h(\cdot)$.

Definition 9 ([6]). *A symmetric $\mathbb{P}^m$-integrable kernel $h : E^m \to \mathbb{R}$ is $\mathbb{P}$-degenerate of order $r-1$ if and only if*

$$\int h(x_1, \ldots, x_m) d\mathbb{P}^{m-r+1}(x_r, \ldots, x_m) = \int h d\mathbb{P}^m$$

holds for any $x_1, \ldots, x_{r-1} \in E$, whereas

$$\int h(x_1, \ldots, x_m) d\mathbb{P}^{m-r}(x_{r+1}, \ldots, x_m),$$

is not a constant function. If h is furthermore $\mathbb{P}^m$-centered, that is, $\mathbb{P}^m f = 0$, we write $h \in L_2^{c,r}(\mathbb{P}^m)$. For notational simplicity, we usually write $L_2^{c,m}(\mathbb{P}^m) = L_2^{c,m}(\mathbb{P})$.

Moreover, $h(\cdot)$ is said to be canonical or completely degenerated if the integral with respect to one variable is equal to zero, i.e.,

$$\int h(x_1, \ldots, x_m) d\mathbb{P}(x_1) = 0 \quad \text{for all} \quad x_2, \ldots, x_m \in E.$$

The fact that the kernel is completely degenerate with the condition $\mathbb{P}^m h^2 < \infty$ is used for the orthogonality of the different elements of the Hoeffding decomposition of the U-statistics.

Definition 10 (Covering number). *The covering number $N_p(\varepsilon, Q, \mathscr{F})$ is defined as the minimal number of balls with radius ε that are needed to cover a class of functions $\mathscr{F}$ in the norm $L_p(Q)$, where Q is the measure on E with finite support.*

We can associate some distances $e_{n,p}$ to the covering numbers, where

$$e_{n,p} = (U_n(|f - g|^p))^{1/p}.$$

In this work, we use the two distances defined afterward

$$e_{n,2}(f,g) = \left(\frac{(n-m)!}{n!} \sum_{0 \leq i_1 < \ldots < i_m \leq n} (f - g)(X_{i_1}, \ldots, X_{i_m})^2 \right)^{1/2}.$$

For decoupled statistics, we also associate covering numbers, well-known as $\widetilde{N}(\epsilon, \mathscr{F}, \widetilde{e}_{n,p})$ and a distance, which can be defined for $p = 2$ as follows:

$$\widetilde{e}_{n,2}(f,g) = n^{1/2} \frac{(n-m)!}{n!} \left[\mathbb{E}_\varepsilon \left(\sum_{0 \leq i_1 < \ldots < i_m \leq n} \varepsilon_{i_1}(f - g)(X_{i_1}, \ldots, X_{i_m}) \right)^2 \right]^{1/2}.$$

Definition 11. *A class $\mathscr{F}$ of measurable functions $E \to \mathbb{R}$ is said to be of VC type (or Vapnik–Chervonenkis type) for an envelope F and admissible characteristic (C, v) (positive constants) such that $C \geq (3\sqrt{e})^v$ and $v \geq 1$, if for all probability measure $\mathbb{Q}$ on $(E, \mathscr{E})$ with $0 < \|F\|_{L_2(Q)} < \infty$ and every $0 < \epsilon < 1$,*

$$N\left(\epsilon \|F\|_{L_2(\mathbb{Q})}, \mathscr{F}, \|\cdot\|_{L_2(\mathbb{Q})} \right) \leq C\epsilon^{-v}.$$

We assume that the class is countable to avoid measurability issues (the noncountable case may be handled similarly by using an outer probability and additional measurability assumptions, see [53]).

Definition 12 (Stochastic equicontinuity, ([54])). *Let $\{Z_n\}$ be a sequence of stochastic processes. Call $\{Z_n\}$ stochastically equicontinuous at t_0 if for each $\delta > 0$, there exists a neighborhood D of t_0 such that*

$$\limsup_n \mathbb{P}\left\{ \sup_D |Z_n(t) - Z_n(t_0)| \right\} < \varepsilon. \tag{6}$$

In the context of the *U*-process $\{U_n\}$, the stochastic equicontinuity at a function $g \in \mathscr{F}$ implies generally that $|U_n(h) - U_n(g)|$ should be uniformly small for all $h(\cdot)$ close enough to $g(\cdot)$, with high probability and for all n large enough.

2.4. Gaussian Chaos Process

Definition 13. *Let H denote a real separable Hilbert space with scalar product $\langle \cdot, \cdot \rangle_H$. We say that a stochastic process $\mathbb{G} = \{\mathbb{G}_\mathbb{P}(h), h \in H\}$ defined in a complete probability space $(E, \mathscr{E}, \mathbb{P})$ is an isonormal Gaussian process (or a Gaussian process on H) if $\mathbb{G}_\mathbb{P}$ is a centered Gaussian family of random variables such that $\mathbb{E}(\mathbb{G}_\mathbb{P}(h)\mathbb{G}_\mathbb{P}(g)) = \langle h, g \rangle_H$ for all $h, g \in H$.*

Define the mapping $h \to \mathbb{G}_\mathbb{P}(h)$. Under the assumption mentioned above, this map is linear and it provides a linear isometry of H onto a closed subspace $L_2(E, \mathscr{E}, \mathbb{P})$ which contains a zero mean Gaussian random variables as its elements. Let $\mathbb{K}_\mathbb{P}$ be the isonormal Gaussian chaos process associated with $\mathbb{G}_\mathbb{P}$ determined by:

$$\mathbb{K}_P\left(h_m^\psi\right) = (m!)^{\frac{1}{2}}\mathfrak{R}_m\left(\mathbb{G}_\mathbb{P}(\psi), \mathbb{E}\psi^2, 0, \ldots, 0\right),$$

where $h_m^\psi(x_1, \ldots, x_m) = \psi(x_1) \cdots \psi(x_m), \psi \in L_2(\mathbb{P})$ and $\mathfrak{R}_m$ is a polynomial defined as a sum of monomials of degree m; ref. [6] give us a simple expression of this polynomial, extracted from Newton's identity given by

$$\sum_{1 \le i_1 < \cdot < i_m \le n} t_{i_1} \cdots t_{i_m} = \mathfrak{R}_m\left(\sum_{i=1}^n t_i, \sum_{i=1}^n t_i^2, \ldots, \sum_{i=1}^n t_i^m\right).$$

Therefore,

$$\sum_{1 \le i_1 < \cdot < i_m \le n} \psi(x_{i_1}) \cdots \psi(x_{i_m}) = \mathfrak{R}_m\left(\sum_{i=1}^n \psi(x_i), \sum_{i=1}^n \psi(x_i)^2, \ldots, \sum_{i=1}^n \psi(x_i)^m\right).$$

Hence, by the continuous mapping theorem, we can see that CLT and LLN give:

$$\left(\binom{n}{m_1}^{\frac{1}{2}} U_n\left(h_{m_1}^{\psi_1}\right), \ldots, \binom{n}{m_r}^{\frac{1}{2}} U_n\left(h_{m_r}^{\psi_r}\right)\right)$$

$$\to \left((m_1!)^{\frac{1}{2}} R_{k_1}\left(\mathbb{G}_\mathbb{P}(\psi_1), \mathbb{E}\psi_1^2, 0, \ldots, 0\right) \ldots, (m_r!)^{\frac{1}{2}} \mathfrak{R}_{k_r}\left(\mathbb{G}_\mathbb{P}(\psi_r), \mathbb{E}\psi_r^2, 0, \ldots, 0\right)\right).$$

Under the linearity of the kernel, we only need to show that:

$$\left\{\binom{n}{m}^{\frac{1}{2}} U_n(f) : f \in \mathscr{F}\right\}$$

$$\xrightarrow{d} \left\{\mathbb{K}_P(f_k) = m!\mathfrak{R}_m\left(\mathbb{G}_\mathbb{P}(\psi), \mathbb{E}\psi^2, 0, \ldots, 0\right) : f_k \in \mathscr{F}\right\} \text{ in } \ell^\infty(\mathscr{F}),$$

to hold the weak convergence. The limit $\mathbb{K}_\mathbb{P}$ is useful in the case of degenerate *U*-statistics and it provides a convergence of all moments, which in turn plays a crucial role because it is due to the hypercontractivity, which makes the uniform integrability better. For a good explanation of $\mathbb{K}_\mathbb{P}$, readers are invited to see ([6] Chapter 4, Section 4.2).

2.5. Technical Assumptions

For our results, we need the following assumptions.

(C.1) (Block-length assumption) For all $q \ge 1$ and $l \ge 1$,

$$\mathbb{E}_\nu\left[\tau^l\right] < \infty, \qquad \mathbb{E}_A\left[\tau^q\right] < \infty;$$

(C.2) (Nonregenerative blocks) For $l \geq 1$, we have

$$
\mathbb{E}_\nu \left[\left(\sum_{i_1=1}^{\mathscr{T}_0} \sum_{i_2=\mathscr{T}_0+1}^{\mathscr{T}_1} \sum_{i_3=\mathscr{T}_1+1}^{\mathscr{T}_2} \cdots \sum_{i_m=\mathscr{T}_{m-1}+1}^{\mathscr{T}_m} |h(X_{i_1},\ldots,X_{i_m})| \right)^l \right] < \infty
$$

and

$$
\mathbb{E}_\nu \left[\left(\sum_{i_1=\mathscr{T}_0+1}^{\mathscr{T}_1} \sum_{i_2=\mathscr{T}_1+1}^{\mathscr{T}_2} \cdots \sum_{i_{m-1}=\mathscr{T}_{m-1}+1}^{\mathscr{T}_m} \sum_{i_m=\mathscr{T}(l_n)+1}^{n} |h(X_{i_1},\ldots,X_{i_{m-1}},X_{i_m})| \right)^l \right] < \infty
$$

(C.3) (Block-sum: moment assumptions) For $l \geq 1$, we have

$$
\mathbb{E}_\nu \left[\left(\sum_{i_1=\mathscr{T}_0+1}^{\mathscr{T}_1} \sum_{i_2=\mathscr{T}_1+1}^{\mathscr{T}_2} \cdots \sum_{i_m=\mathscr{T}_{m-1}+1}^{\mathscr{T}_m} |h(X_{i_1},\ldots,X_{i_m})| \right)^l \right] < \infty,
$$

and

$$
\mathbb{E}_A \left[\left(\sum_{\mathscr{T}_0+1 \leq i_1 \leq \ldots \leq i_m \leq \mathscr{T}_1} h(X_{i_1},\ldots,X_{i_m}) \right)^l \right] < \infty;
$$

(C.4) For $l \geq 1$, we have

$$
\mathbb{E}_\nu \left[\left(\sum_{i_1=\mathscr{T}_0+1}^{\mathscr{T}_1} \sum_{i_2=\mathscr{T}_1+1}^{\mathscr{T}_2} \underbrace{\sum_{i_k=\mathscr{T}_k+1}^{\mathscr{T}_{k+1}} \cdots \sum_{i_k=\mathscr{T}_k+1}^{\mathscr{T}_{k+1}}}_{u \text{ times}} \sum_{i_{k+u}=\mathscr{T}_{k+u}+1}^{\mathscr{T}_{k+u+1}} \cdots \sum_{i_m=\mathscr{T}_{m-1}+1}^{\mathscr{T}_m} \right.
$$

$$
\left. |h(X_{i_1},\underbrace{X_{i_k},\ldots,X_{i_k}}_{u \text{ times}},X_{i_{k+u}},\ldots,X_{i_m})| \right)^l \right] < \infty;
$$

(C.5) (Nondegeneracy.) We suppose also that

$$
\mathbb{E}_A \left[\left(\sum_{i=\mathscr{T}_0+1}^{\mathscr{T}_1} h_1(X_i) \right)^2 \right] > 0.
$$

Remark 1 (Moment assumptions). *In practice, we recall that block-moment assumptions for the split Markov chain can be generally checked by establishing drift conditions of Lyapunov's type for the original chain; see Chapter 11 in [55,56], as well as All these moment conditions are discussed in detail in ([57], Chapters 11 and 17). There is a key condition in the proof of ergodic theorems in the Markovian context, which is the fact that $\mathbb{E}_\mathbf{A}(\tau_0) < \infty$, for any $\mathbf{A}$ that is a set in $\mathscr{E}$, such that $\psi(\mathbf{A}) > 0$. In fact, when there is a finite invariant measure and an atom $\mathbf{A}$, then this condition is readily found. We also refer to [58] for an explicit check of such conditions on several important examples and to §4.1.2 of [59] for sufficient conditions expressed in terms of a uniform return rate to small sets. Finally, as discussed in Chapter 8 of [60], similar conditions can be expressed in potential kernels. Observe that, in the positive recurrent case, the assumptions of (C.1) are not independent when $\nu = \mu$: from the basic renewal theory, one has $\mathbb{P}_\mu(\tau = k) = (\mathbb{E}_A[\tau])^{-1}\mathbb{P}_A(\tau \geq k)$ for all $k \geq 1$. Hence, conditions $\mathbb{E}_\mu[\tau^l] < \infty$ and $\mathbb{E}_A[\tau^{l+1}] < \infty$ are equivalent.*

3. Preliminary Results

A significant issue was detected in recovering the estimation of our parameter of interest using the U-process. The given shape of this parameter is as follows:

$$\mu(h) = \int_{x_1 \in E} \cdots \int_{x_k \in E} h(x_1, \ldots, x_k) \mu(dx_1) \ldots \mu(dx_k),$$

where $h : E^m \to \mathbb{R}$ is a kernel function. The estimation of this parameter should be possible using the U-statistics of the form:

$$U_n(h) = \binom{n}{m}^{-1} \sum_{(i_1, \ldots, i_m) \in I_n^m} h(X_{i_1}, \ldots, X_{i_m}), \text{ for } n \geq m, \tag{7}$$

As the parameter of interest is defined and based on Kac's theorem for the occupation measure, $\mu(h)$ in the regeneration setup can be written as follows:

$$\mu(h) = \frac{1}{(\mathbb{E}_A(\tau))^m} \mathbb{E}_A \left(\sum_{i_1 = \mathscr{T}_0 + 1}^{\mathscr{T}_1} \sum_{i_2 = \mathscr{T}_1 + 1}^{\mathscr{T}_2} \cdots \sum_{i_m = \mathscr{T}_{(m-1)} + 1}^{\mathscr{T}_m} h(X_{i_1}, \ldots, X_{i_m}) \right). \tag{8}$$

In the Markovian context and since the variables are not independent, the approximation related to the i.i.d. blocks and the regenerative case is introduced below:

Definition 14 (Regenerative kernel). *Let $h : E^m \to \mathbb{R}$ a kernel. We define the regenerative kernel $\omega_h : \mathbb{T}^m \to \mathbb{R}$ as follows:*

$$\omega_h((x_{11}, \ldots, x_{1n_1}), \ldots, (x_{k1}, \ldots, x_{kn_k})) = \sum_{i_1 = 1}^{n_1} \cdots \sum_{i_k = 1}^{n_k} h(x_{1i_1}, \ldots, x_{ki_k}). \tag{9}$$

It is not necessary that the kernel $\omega_h(\cdot)$ be symmetric, as soon as $h(\cdot)$. In fact, we can use the symmetrization of $S_m \omega_h$ in the following way

$$(S_m \, \omega_h) = (m!)^{-1} \sum_{\sigma} \sum_{\sigma(1)=1}^{n_1} \cdots \sum_{\sigma(m)=1}^{n_k} h\left(x_{\sigma(1)}, \ldots, x_{\sigma(m)} \right), \tag{10}$$

where the first sum is over all permutations $\sigma = \{i_1, \ldots, i_m\}$ of $\{1, \ldots, m\}$. Next, we consider the U-statistic formed by the regenerative data.

Definition 15 (Regenerative U-statistic). *Let $h : E^m \to \mathbb{R}$ a kernel such that $\mu(|h|) < \infty$ and set $\widetilde{h}(\cdot) = h(\cdot) - \mu(h)$. The regenerative U-statistic associated with the sequence of regenerative blocks $\{\mathbf{B}_j\}_{j=1}^L$, generated by the Markov chain is given by*

$$R_{l_n}(h) = \binom{l_n - 1}{m}^{-1} \sum_{(i_1, \ldots, i_m) \in I_{l_n - 1}^m} \omega_{\widetilde{h}}(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m}). \tag{11}$$

Hence, $R_{l_n}(h)$ is a standard U-statistic with mean zero.

Proposition 1. *Let us define*

$$W_n(h) = U_n(h) - \mu(h) - \binom{l_n - 1}{m} \binom{n}{m}^{-1} R_{l_n}(h). \tag{12}$$

Then, under conditions (C.1), (C.2), (C.3) and (C.4), we have the following stochastic convergences:

$$W_n(h) \rightarrow 0, \quad \mathbb{P}_\nu - a.s.$$

Before stating the weak convergence in the next theorem, we define the corresponding *U*-processes related to the *U*-statistic U_n and the regenerative *U*-statistic R_L, respectively:

$$Z_n := \binom{n}{m}^{1/2} [U_n - \mu(h)], \tag{13}$$

$$T_{l_n} := \binom{l_n}{m}^{1/2} [R_{l_n} - \mathbb{E}(R_{l_n})]. \tag{14}$$

Theorem 1. *Let $(X_n)_n$ be a positive recurrent Harris Markov chain, with an accessible atom* **A**, *X_n satisfies the conditions (C.1) and (C.2) (moments assumptions), (C.3), (C.4), (C.5) and, for a fixed $\gamma > 0$, $\mathbb{E}(\tau)^{2+\gamma} < \infty$. Let $\mathscr{F}$ be a uniform bounded class of functions with an envelope H square-integrable such that:*

$$\int_0^\infty (\log N(\varepsilon, \mathscr{F}, e_{n,2}))^{m/2} d\varepsilon < \infty.$$

Then, the process Z_n converges weakly in probability under $\mathbb{P}_\nu$ to a Gaussian process $\mathbb{G}_\mathbb{P}$ indexed by $\mathscr{F}$ whose sample paths are bounded and uniformly continuous with respect to the metric $L_2(\mu)$.

The Bootstrapped U-Processes

Trying to facilitate the bootstrap technique, we write the detailed steps of the regenerative block construction and the weighted bootstrap method in Algorithm 1:

Algorithm 1 Regenerative block and weighted bootstrap construction.

---

1. Identify the number of visits $l_n = \sum_{i=0}^n \mathbb{1}_{X_i \in A}$ to the atom A.
2. Divide the sample $X^{(n)} = (X_1, \ldots, X_n^{(n)})$ into $(l_n + 1)$ regenerative blocks $\mathbf{B}_0, \ldots, \mathbf{B}_{l_n-1}, \mathbf{B}_{l_n}^{(n)} \in \mathbb{T}$, each block B_i with a length $l(\mathbf{B}_i) \equiv \tau_i$.
3. Drop the first and the last blocks if $\tau_{l_n} < n$ to avoid bias.
4. Let $\boldsymbol{\xi} = (\xi_{i,l_n}, i = 1, \ldots, n)$ be a triangular array of random variables. Define the weighted bootstrap empirical measure from the data:

$$\mathbb{P}_n^* = \frac{1}{l_n} \sum_{i=1}^n \xi_{i,l_n} \delta_{\mathbf{B}_i}.$$

In what follows, we denote by $\mathbb{P}^*$ and $\mathbb{E}^*$, respectively, the conditional probability and the conditional expectation given the sample $\{X_1, \ldots, X_n\}$. The same notation is used for the sample $\{\mathbf{B}_1, \ldots, \mathbf{B}_{L_n}\}$. Define the bootstrapped *U*-statistic as

$$U_n^*(h) = \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m) \in I_n^m} \xi_{i_1,n} \ldots \xi_{i_m,n} h(X_{i_1}, \ldots, X_{i_m}) \tag{15}$$

and the regenerative bootstrapping

$$R_{l_n}^*(h) = \binom{l_n}{m}^{-1} \sum_{(i_1,\ldots,i_m) \in I_{l_n}^m} \xi_{i_1,l_n} \ldots \xi_{i_m,l_n} \omega_h(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m}). \tag{16}$$

and the *U*-processes are:

$$Z_n^* := \binom{n}{m}^{1/2} [U_n^*(h) - U_n(h)]_{h \in \mathscr{F}}$$

$$= \binom{n}{m}^{-1/2} \sum_{(i_1,\dots,i_m)\in I_n^m} (\xi_{i_1,l_n} - 1)\dots(\xi_{i_m,l_n} - 1)h(X_{i_1},\dots,X_{i_m}). \qquad (17)$$

and

$$
\begin{aligned}
T_{l_n}^* &:= \binom{l_n}{m}^{1/2}\left[R_{l_n}^*(h) - \mathbb{E}(R_{l_n}^*(h))\right]_{h\in\mathscr{F}} \\
&= \binom{l_n}{m}^{-1/2} \sum_{(i_1,\dots,i_m)\in I_{l_n}^m} (\xi_{i_1,l_n} - 1)\dots(\xi_{i_m,l_n} - 1)\omega_h(\mathbf{B}_{i_1},\dots,\mathbf{B}_{i_m}). \qquad (18)
\end{aligned}
$$

Given Δ_n, a real-valued function, defined on the product probability space, we say that Δ_n is of an order $o_{\mathbb{P}_\xi}^o(1)$ in $\mathbb{P}_\nu^o$-probability if for any $\varepsilon, \delta > 0$

$$\mathbb{P}_\nu^o\left\{\mathbb{P}_{\xi|X}^o(|\Delta_n| > \varepsilon) > \delta\right\} \longrightarrow 0 \quad \text{as } n \to \infty$$

and that Δ_n is of an order $O_{\mathbb{P}_\xi}^o(1)$ in $\mathbb{P}_\nu^o$-probability if for any $\delta > 0$, there exists a $0 < M < \infty$ such that

$$\mathbb{P}_\nu^o\left\{\mathbb{P}_{\xi|X}^o(|\Delta_n| \geq M) > \delta\right\} \longrightarrow 0 \quad \text{as } n \to \infty$$

We must comment here that the bootstrap works in probability if

$$d_{BL}(T_{l_n}^*, T_{l_n}) \to 0 \text{ in probability,}$$

where

$$d_{BL}(T_{l_n}^*, T_{l_n}) = \sup_{g\in BL(l^\infty(\mathscr{F}))}\left|\mathbb{E}g\left(T_{l_n}^*\right)^\circ - \mathbb{E}g(T_{l_n})\right|,$$

and

$$BL(l^\infty(\mathscr{F})) := \{g : l^\infty(\mathscr{F}) \to \mathbb{R}, |g(x) - g(y)| \leq \|x - y\|_{\mathscr{F}}, \|g\|_\infty \leq 1\},$$

and $g\left(T_{l_n}^*\right)^\circ$ is the measurable envelope of $g\left(T_{l_n}^*\right)$. In addition, for any measurable random elements, Y_n and Y, the convergence in law of Y_n to Y is in the sense of Hoffman–Jorgensen, which is defined as

$$\mathbb{E}g(Y_n)^\circ \to \mathbb{E}g(Y),$$

for g bounded and continuous. This weak convergence is metrizable by Theorem A1 in Appendix A.

Proposition 2. *Suppose that the bootstrap weights $(\xi_1,\dots,\xi_n)$ satisfy Assumptions (A1)–(A4). Let*

$$W_n^*(h) := U_n^*(h) - \binom{l_n - 1}{m}\binom{n}{m}^{-1}R_{l_n}^*(h). \qquad (19)$$

Then, we have

$$W_n^*(h) \quad \to \quad 0, \quad \mathbb{P}_\nu \times \mathbb{P}_\xi - a.s.$$

The proof of Proposition 2 is postponed until Section 7.
Now, in the following lemma, there are some instrumental results needed later.

Lemma 1. *Let $(X_n)_n$ be a Markov chain defined in Section 2.1. Define $p := \mathbb{P}(X_0 \in A) = \alpha^{-1}$. Then, for any initial probability ν, we have:*

(i)   *For some $\eta > 0$ and $C > 0$:*

$$\left| \frac{\mathbb{E}_\nu(l_n)}{np} - 1 \right| \leq \frac{C}{n} \quad and \quad \sqrt{n}\left(\frac{l_n}{np} - 1 \right) \to N(0, \eta^2). \tag{20}$$

(ii) $\frac{n*}{n} \to 1$ *in $\mathbb{P}_\nu \times \mathbb{P}_\xi$-probability .*

(iii)   *Let X_i be a sequence of random variables. If*

$$T_n = \frac{1}{n} \sum_{i=1}^{n} X_i \to C \ \ a.s.,$$

then for any integer t_n valued sequence of random variables,

$$\frac{1}{t_n} \sum_{i=1}^{t_n} X_i \to C \ \ in \ \mathbb{P}_\nu\text{-probability.}$$

The proof of Lemma 1 is postponed until Section 7.

4. Weighted Bootstrap Weak Convergence

In this section, we extend some existing results concerning the multiplier U-process to prove the bootstrap uniform weak convergence. Most of these results can be found in [46], generalizing the empirical process work of [38] in the i.i.d. setting. The weak convergence is proved for degenerate U-processes, as we mentioned before, and under the weighted regenerative bootstrap schemes described in Algorithm 1. Before stating the weak convergence theorem, we recall the following important results. The next theorem, proved in [46], is a sharp multiplier inequality, which is essential in the study of the multiplier U-process. These results are based on the decoupling symmetrized U-process, a basic framework of U-statistics. In [47], the author solved these problems for the empirical process settings in the Markov setting (multinomial bootstrap), which we generalize to the U-process by considering more general weights, i.e., the exchangeable weighted bootstrap.

Theorem 2 ([46]). *Let $(\xi_1, \ldots, \xi_n)$ be a random vector independent of $(Y_1, \ldots, Y_n)$. Then, there exists some measurable function $\psi_n : \mathbb{R}_{\geq 0}^m \to \mathbb{R}_{\geq 0}$ such that the expected supremum of the decoupled (Here "decoupled" refers to the fact that $\{Y_i^{(k)}\}, k \in \mathbb{N}$ are independent copies of $\{Y_i\}$, and $\{\epsilon_i^{(k)}\}, k \in \mathbb{N}$ are independent copies of the Rademacher sequence $\{\epsilon_i\}$.) U-processes*

$$\mathbb{E} \left\| \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \epsilon_{i_1}^{(1)} \cdots \epsilon_{i_m}^{(m)} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)} \right) \right\|_{\mathscr{F}} \leq \psi_n(\ell_1, \ldots, \ell_m),$$

for all $1 \leq \ell_1, \ldots, \ell_m \leq n$, consequently,

$$\mathbb{E} \left\| \sum_{1 \leq i_1, \ldots, i_m \leq l_{n-1}} \xi_{i_1} \cdots \xi_{i_m} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)} \right) \right\|_{\mathscr{F}}$$

$$\leq K \int_{\mathbb{R}_{\geq 0}^m} \mathbb{E} \psi_n \left(\sum_{i=1}^{l_{n-1}} \mathbb{1}_{|\xi_i| > t_1}, \ldots, \sum_{i=1}^{l_{n-1}} \mathbb{1}_{|\xi_i| > t_m} \right) dt_1 \cdots dt_m.$$

Furthermore, if there exists a concave and nondecreasing function $\bar{\psi}_n : \mathbb{R} \to \mathbb{R}$ such that $\psi_n(\ell_1, \ldots, \ell_m) = \bar{\psi}_n \left(\prod_{k=1}^m \ell_k \right)$, then

$$\mathbb{E} \left\| \sum_{1 \leq i_1, \ldots, i_m \leq n} \xi_{i_1} \cdots \xi_{i_m} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)} \right) \right\|_{\mathscr{F}}$$

$$\leq \quad K \int_{\mathbb{R}_{\geq 0}^m} \bar{\psi}_n \left(\sum_{1 \leq i_1, \ldots, i_m \leq n} \prod_{k=1}^m \mathrm{P} \left(|\xi_{i_k}| > t_k \right)^{1/m} \right) dt_1 \cdots dt_m.$$

Here, $K > 0$ is a constant depending on m only and can be taken as $K = 2^{2m} \prod_{k=2}^m (k^k - 1)$ for $m \geq 2$.

Lemma 2 ([46])**.** *Let $\{ \mathscr{F}_{(\ell_1, \ldots, \ell_m), n} : 1 \leq \ell_1, \ldots, \ell_m \leq n, n \in \mathbb{N} \}$ be function classes such that $\mathscr{F}_{(\ell_1, \ldots, \ell_m), n} \supset \mathscr{F}_{(n, \ldots, n), n}$ for all $1 \leq \ell_1, \ldots, \ell_m \leq n$. Suppose that the ξ_i's have the same marginal distributions with $\|\xi_1\|_{2m,1} < \infty$. Suppose that there exists some bounded measurable function $a : \mathbb{R}_{\geq 0}^{m(n)} \to \mathbb{R}_{\geq 0}$ with $a(\ell_1, \ldots, \ell_m) \to 0$ as $\ell_1 \wedge \ldots \wedge \ell_m \to \infty$, such that the expected supremum of the decoupled U-processes satisfies*

$$\mathbb{E} \left\| \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \epsilon_{i_1}^{(1)} \cdots \epsilon_{i_m}^{(m)} f \left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^m \right) \right\|_{\mathscr{F}_{(\ell_1, \ldots, \ell_m), n}}$$

$$\leq a(\ell_1, \ldots, \ell_m) \left(\prod_{k=1}^m \ell_k \right)^{1/2},$$

for all $1 \leq \ell_1, \ldots, \ell_m \leq n$. Then,

$$n^{-m/2} \mathbb{E} \left\| \sum_{1 \leq i_1, \ldots, i_m \leq n} \xi_{i_1} \cdots \xi_{i_{m(n)}} f \left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)} \right) \right\|_{\mathscr{F}_{(n, \ldots, n), n}} \to 0, \quad n \to \infty.$$

The main result of this paper is represented in the following theorem. It is worth noting here that it is not easy to prove the stochastic equicontinuity in the present setting as explained in the introduction.

Definition 16 (Permissible classes of function)**.** *Let $(E, \mathscr{E}, \mathbb{P})$ be a measurable space ($\mathscr{E}$ a Borel σ-field on E). Let $\mathscr{F}$ be a class of functions indexed by a parameter x that belongs to a set E. $\mathscr{F}$ is called permissible if it can be indexed by a E such that:*

- *There exists a function $g(x, f) = f(x)$ defined from $S \times \mathscr{F}$ to $\mathbb{R}$ in such a way that this function is $\mathscr{L} \otimes \mathscr{B}(\mathscr{F})$ measurable function, where $\mathscr{B}(\mathscr{F})$ is the Borel σ-algebra generated by the metric on $\mathscr{F}$.*
- *E is a Suslin measurable space whose mean E is an analytic subset of a compact metric space E from which it inherits its metric and Borel σ-field.*

Theorem 3. *Suppose Assumptions (A1) to (A4), and Conditions (C.1)–(C.5) hold. Let $\mathscr{F} \subset L_2^{c,m}(P)$ be permissible and admit a P^m-square integrable envelope F such that*

$$\int_0^1 \left(\sup_Q \log \mathscr{N} \left(\epsilon \|F\|_{L_2(Q)}, \mathscr{F}, L_2(Q) \right) \right)^{m/2} d\epsilon < \infty,$$

where the supremum is taken over all discrete probability measures. Then,

$$\sup_{\psi \in \mathrm{BL}} \left| \mathbb{E}_\zeta \psi \left(Z_n^*(h) \right) - \mathbb{E} \psi (c \cdot \mathbb{K}_\mathbb{P}) \right| \to_{\mathbb{P}_\nu} 0,$$

where c is the constant in (A3), and the convergence in probability $\to_{\mathbb{P}_\nu}$ is with respect to the outer probability of $\mathbb{P}^\infty$ defined on $(E^\infty, \mathscr{E}^\infty)$.

The proof of Theorem 3 is postponed until Section 7.

4.1. Bootstrap Weights Examples

Let $(\xi_1, \cdots, \xi_n)$ be a class of real random variables satisfying Assumptions *(A1)–(A4)*. We give some examples of bootstrap weights; for instance, refer to [38,61] for more explanations.

4.1.1. Bayesian Resampling Scheme

In this case, $(\xi_1, \ldots, \xi_n)$ are positive i.i.d. random variables with mean μ and finite variance σ^2. The weights satisfy $\|\xi_1\|_{2,1} < \infty$, and we define

$$\overline{\xi}_n = \sum_{i=1}^{n} \xi_i.$$

The Bayesian bootstrapped weight can be defined as:

$$\xi_{ni} = \xi_i / \overline{\xi}_n,$$

satisfying

$$\|\xi_{n1}\|_{2,1} = \int_0^{\infty} \sqrt{\mathbb{P}_{\xi}(\xi_{n1} \geq u)} \, du.$$

For $\xi_{ni} \sim Exponential(1)$ or $\xi_{ni} \sim Gamma(4,1)$, the Bayesian weights are distributionally equivalent with Dirichlet weights. For the value of c^2, we have:

$$\frac{1}{n} \sum_{i=1}^{n} (\xi_{ni} - 1)^2 \to \frac{\sigma^2}{\mu^2} := c^2, \quad n \to \infty.$$

4.1.2. Efron's Resampling Scheme

For Efron's bootstrap, we have

$$(\xi_1, \ldots, \xi_n) \sim \text{Multinomial}(n; n^{-1}, \ldots, n^{-1}).$$

Condition *(A1)* follows directly. Condition *(A3)* follows from ([37] Lemma 4.1), and Condition *(A2)* is detailed in [43].

4.1.3. The Delete h-Jackknife

In [62], the authors permute deterministic weights w_n, where

$$w_n = \left\{ \frac{n}{n-h}, \ldots, \frac{n}{n-h}, 0, \ldots, 0 \right\} \quad \text{with} \quad \sum_{i=1}^{n} w_{ni} = n$$

in order to build new bootstrap weights, and they defined the new weights $\xi_{nj} := w_{nR_n(j)}$ where $R_n(\cdot)$ is a random permutation uniformly distributed over $\{1, \ldots, n\}$. These weights are called the delete h-Jackknife. In order to achieve Assumption *(A3)*, we must assume that $h/n \to \alpha \in (0,1)$, as $c^2 = h/(n-h)$ and $c > 0$.

4.1.4. The Multivariate Hypergeometric Resampling Scheme

As its name indicates, the bootstrap weights of this scheme follow the multivariate hypergeometric distribution with density:

$$P(\xi_{n1} = \varepsilon_1, \ldots, \xi_{nn} = \varepsilon_n) = \frac{\binom{K}{\varepsilon_1} \cdots \binom{K}{\varepsilon_n}}{\binom{nK}{n}},$$

where K is a positive integer. Assumption *(A3)* is satisfied with $c^2 = (K-1)/K$.

Remark 2. *As was pointed out in [38], the preceding mentioned bootstraps are "smoother" in some way than the multinomial bootstrap because they place some (random) weight on all elements*

in the sample, whereas the multinomial bootstrap applies the positive weight at a proportion of about $1 - (1 - n^{-1})^n \to 1 - e^{-1} = 0.6322$ of each element of the sample, on average. Notice that when $\omega_i \sim Gamma(4, 1)$, the ξ_{ni}/n are equivalent to four spacings from a sample of $4n - 1$ Uniform $(0, 1)$ random variables. In [63,64], it was noticed that in addition to being four times more expensive to implement, the choice of four spacings depends on the functional of interest and is not universal.

Remark 3. *It is noteworthy that choosing the bootstrap weights ξ_{ni} properly implies a smaller limit variance, that is, c^2 is smaller than 1. A typical example is the multivariate hypergeometric bootstrap ([38] Example 3.4) and the subsample bootstrap, ([65] Remark 2.2-(3)). A thorough treatment of the weight selection is undoubtedly outside the scope of the current work; for review, we refer the readers to [66].*

Remark 4. *In the present paper, we considered a renewal type of bootstrap for atomic Markov chains under minimal moment conditions on renewal times. The atomic Markov chains assumption can be dropped by mimicking the ideas of [50,51] by introducing an artificial atom and deriving the bootstrap procedure that applies to nonatomic Markov chains. Precisely, in the case of a general irreducible chain X with a transition kernel $\Pi(x, dy)$ satisfying a minorization condition:*

$$\forall x \in S, \Pi(x, dy) \geqslant \delta \psi(dy),$$

for an accessible measurable set S, a probability measure ψ and $\delta \in]0, 1[$ (note that such a minorization condition always holds for Π or an iterate when the chain is irreducible), an atomic extension (X, Y) of the chain may be explicitly constructed by the Nummelin splitting technique (see [49]) from the parameters (S, δ, ψ) and the transition probability Π, see for instance [47,67]. From a practical viewpoint, the size of the first block may be large compared to the size n of the whole trajectory, for instance, in the case where the expected return time to the (pseudo-)atom when starting with the initial probability distribution is large. The effective sample size for constructing the data blocks and the corresponding statistic is then dramatically reduced. However, in [68], some simulations were given together with examples including content-dependent storage systems and general AR models supporting the method discussed in this work.

5. Applications

Example 1 (Symmetry test). *This example gives an application for the bootstrap U-statistics, inspired by the goodness-of-fit tests in [69], where they considered the symmetry test for the distribution of X_t. Let $(X_t)_{t \in \mathbb{N}}$ be a stationary mixing process with $f_X(\cdot)$ the Lebesgue density. We test the hypothesis:*

$$\begin{cases} H_0: & f_X(u) & = f_X(-u), & \text{almost every were,} \\ H_1: & f_X(u) & \neq f_X(-u) & \text{on a set of positive measure.} \end{cases} \quad (21)$$

The estimator of $f_X(u)$ is:

$$\hat{f}_X(u) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{u - X_i}{h_n}\right),$$

where $K(\cdot)$ is a kernel function and $h_n > 0$ is a smoothing parameter or the bandwidth. An appropriate estimator of the integrated squared difference represent the symmetry test:

$$I = \int_{\mathbb{R}} (f_X(u) - f_X(-u))^2 du.$$

According to [69], I can be estimated by

$$\hat{I}_n := \frac{4}{n^2 h_n} \sum_{1 \leq i < j \leq n} \Phi_n(X_i, X_j),$$

where $\Phi_n(X_i, X_j) = K_{X_i, X_j} - K_{X_i, -X_j}$ with $K_{X_i, Y_j} = K\left(\frac{X_i - Y_j}{h_n}\right)$, for $Y_j \in \{X_j, -X_j\}$. Clearly, $\hat{I}_n$ is a degenerate U-statistic with kernel varying with the sample size n. Thus, the stationary bootstrap test,

$$\hat{I}_n^* := \frac{4}{n^2 h_n} \sum_{1 \leq i < j \leq n} \Phi_n\left(X_i^*, X_j^*\right),$$

can be shown to have the same limit as $\hat{I}_n$.

Example 2 (**Kendall's tau**). *The covariance matrix quantifies the linear dependency in a random vector. The rank correlation is another measure of the nonlinear dependency in a random vector. Two generic vectors $y = (y_1, y_2)$ and $z = (z_1, z_2)$ in $\mathbb{R}^2$ are said to be concordant if $(y_1 - z_1)(y_2 - z_2) > 0$. For $m, k = 1, \ldots, p$, define*

$$\tau_{mk} = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \mathbf{1}\left\{\left(X_{im} - X_{jm}\right)\left(X_{ik} - X_{jk}\right) > 0\right\}.$$

Then, Kendall's tau rank correlation coefficient matrix $T = \{\tau_{mk}\}_{m,k=1}^p$ is a matrix-valued U-statistic with a bounded kernel. It is clear that τ_{mk} quantifies the monotonic dependency between (X_{1m}, X_{1k}) and (X_{2m}, X_{2k}) and it is an unbiased estimator of

$$\mathbb{P}((X_{1m} - X_{2m})(X_{1k} - X_{2k}) > 0),$$

that is, the probability that (X_{1m}, X_{1k}) and (X_{2m}, X_{2k}) are concordant.

Example 3 (**Test of independence**). *In [2] the author introduced the parameter*

$$\triangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D^2(y_1, y_2) dF(y_1, y_2),$$

where $D(y_1, y_2) = F(y_1, y_2) - F(y_1, \infty)F(\infty, y_2)$ and $F(\cdot, \cdot)$ is the distribution function of Y_1 and Y_2. The parameter $\triangle$ has the property that $\triangle = 0$ if and only if Y_1 and Y_2 are independent. From [8], an alternative expression for $\triangle$ can be developed by introducing the functions

$$\psi(y_1, y_2, y_3) = \begin{cases} 1 & \text{if } y_2 \leq y_1 < y_3 \\ 0 & \text{if } y_1 < y_2, y_3 \text{ or } y_1 \geq y_2, y_3 \\ -1 & \text{if } y_3 \leq y_1 < y_2 \end{cases}$$

and

$$\varphi(y_{1,1}, y_{1,2}, \ldots, y_{5,1}, y_{5,2}) = \frac{1}{4} \psi(y_{1,1}, y_{1,2}, y_{1,3}) \psi(y_{1,1}, y_{1,4}, y_{1,5}) \psi(y_{1,2}, y_{2,2}, y_{3,2}) \psi(y_{1,2}, y_{4,2}, y_{5,2}).$$

We have

$$\triangle = \int \ldots \int \varphi(y_{1,1}, y_{1,2}, \ldots, y_{5,1}, y_{5,2}) dF(y_{1,1}, y_{1,2}) \ldots dF(y_{1,5}, y_{2,5}).$$

The corresponding U-statistics may be used to test the independence.

6. Conclusions

The present paper was concerned with the randomly weighted bootstrap of the U-process in a Markov framework. A large number of bootstrap resampling schemes emerged as special cases of our setting, in particular, the multinomial bootstrap, which is the best-known bootstrap scheme introduced by [26]. One of the main tools was the approximation of the Markov U-process by the corresponding regenerative one. We looked to mimic this result in Proposition 2, in order to approximate the weighted-bootstrap U-process U_n^* to the regenerative weighted-bootstrap U-process $R_{l_n}^*$. Other technical arguments were given in Lemma 1 extended from the work of [47]. These intricate tools were used to reach the full independence of regenerative block variables by proving that a deterministic one could

substitute the random size of blocks, which was the main problem for the extension of the bootstrap results to the Markov framework. After a lengthy proof to arrive at independence, we used the results of [46]. All the above steps led us to prove the weak convergence of the regenerative-block weighted-bootstrap *U*-process, which implied the weak convergence of the weighted-bootstrap *U*-process. It will be of interest to consider the extension of the paper to the semi-Markov setting. A more delicate problem is to consider the setting of incomplete data such as censored cases or missing data. To the best of our knowledge, this problem has not been considered, even for the original sample (without bootstrap) in the Markov framework. It would be interesting to extend our work to the case of the local stationary process, which requires nontrivial mathematics; this would go well beyond the scope of the present paper.

7. Mathematical Development

This section is devoted to the proof of our results. The previously defined notations continue to be used in what follows.

Proof of Proposition 2. We have

$$
U_n^*(h) - \binom{l_n - 1}{m}\binom{n}{m}^{-1} R_{l_n}^*(h)
$$

$$
= \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_n^m} \xi_{i_1,n^*}\cdots\xi_{i_m,n^*} h(X_{i_1},\ldots,X_{i_m})
$$

$$
- \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_{l_n}^m} \xi_{i_1,l_n}\cdots\xi_{i_m,l_n}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})
$$

$$
= \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_n^m} \xi_{i_1,n^*}\cdots\xi_{i_m,n^*} h(X_{i_1},\ldots,X_{i_m})
$$

$$
- \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_{l_n}^m} \xi_{i_1,l_n}\cdots\xi_{i_m,l_n}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})
$$

$$
+ \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_n^m} h(X_{i_1},\ldots,X_{i_m}) - \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_n^m} h(X_{i_1},\ldots,X_{i_m})
$$

$$
+ \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_{l_n}^m} \{\xi_{i_1,l_n}\cdots\xi_{i_m,l_n} - 1\}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})
$$

$$
- \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_{l_n}^m} \{\xi_{i_1,l_n}\cdots\xi_{i_m,l_n} - 1\}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})
$$

$$
= \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_n^m} \{\xi_{i_1,n^*}\cdots\xi_{i_m,n^*} - 1\} h(X_{i_1},\ldots,X_{i_m})
$$

$$
- \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_{l_n}^m} \{\xi_{i_1,l_n}\cdots\xi_{i_m,l_n} - 1\}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})
$$

$$
+ \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_{l_n}^m} \omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})
$$

$$
- \binom{n}{m}^{-1} \sum_{(i_1,\ldots,i_m)\in I_n^m} h(X_{i_1},\ldots,X_{i_m}).
$$

Given $J \subseteq \{1,\ldots,m\}$ ($J = \varnothing$ is not excluded), and $\mathbf{i} = (i_1,\ldots,i_m) \in \{1,\ldots,n\}^m$, we set $\mathbf{i}_J$ to be the point of $\{1,\ldots,n\}^{|J|}$ obtained from i by deleting the coordinates in the places not

in J (e.g., if $\mathbf{i} = (3, 4, 2, 1)$, then $\mathbf{i}_{\{1,3\}} = (3, 2)$). Furthermore, $\sum_{i_J}$ indicates the sum over $1 \leq i_j \leq n, j \in J$; for instance, if $m = 4$ and $J = \{1, 3\}$, then

$$\sum_{\mathbf{i}_J} h_{\mathbf{i}} = \sum_{\mathbf{i}_{\{1,3\}}} h_{i_1, i_2, i_3, i_4} = \sum_{1 \leq i_1, i_3 \leq n} h_{i_1, i_2, i_3, i_4}\left(X_{i_1}^{(1)}, \ldots, X_{i_4}^{(4)}\right).$$

By convention, $\sum_{\mathbf{i}} a = a$. Notice that

$$\mathbb{E}\binom{n}{m}^{-1} \sum_{(i_1, \ldots, i_m) \in I_n^m} \{\xi_{i_1,n} \ldots \xi_{i_m,n} - 1\} h(X_{i_1}, \ldots, X_{i_m})$$

$$= \binom{n}{m}^{-1} \sum_{(i_1, \ldots, i_m) \in I_n^m} \mathbb{E}\{\xi_{i_1,n} \ldots \xi_{i_m,n} - 1\} \mathbb{E} h(X_{i_1}, \ldots, X_{i_m})$$

$$= \binom{n}{m}^{-1} \sum_{\mathbf{i}_{\{1, \ldots, m-1\}}} \sum_{j=1}^{m} \mathbb{E}\left\{ \prod_{k=1, k \neq j}^{m} \xi_{i_k,n} \mathbb{E}\left\{ \sum_{i_j=1}^{n} (\xi_{i_j,n} - 1) \mid \prod_{k=1, k \neq j}^{m} \xi_{i_k,n} \right\} \right\}$$

$$\times \mathbb{E} h(X_{i_1}, \ldots, X_{i_m})$$

$$= 0.$$

In a similar way, we have

$$\mathbb{E}\binom{n}{m}^{-1} \sum_{(i_1, \ldots, i_m) \in I_{l_n}^m} \{\xi_{i_1, l_n} \ldots \xi_{i_m, l_n} - 1\} \omega_{\widetilde{h}}(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m}) = 0.$$

Making use of Proposition 1 and the law of large numbers, we infer that

$$U_n^*(h) - \binom{l_n - 1}{m} \binom{n}{m}^{-1} R_{l_n}^*(h) = 0, \quad a.s.$$

Hence, the proof is completed. $\qquad \square$

Proof of Lemma 1. The proof of part *i*) and part *iii*) follows from ([47] Lemma 3.1 and Lemma 3.2). In order to prove *ii*), we need to show that, for every $\epsilon > 0$,

$$\mathbb{P}_\nu \times \mathbb{P}_{\xi|X}\left(\left|\frac{n*}{n} - 1\right| > \epsilon\right) \to 0, \tag{22}$$

which follows if, conditioned on the sample,

$$\mathbb{P}_{\xi|X}\left(\left|\frac{n*}{n} - 1\right| > \epsilon\right) \to 0. \tag{23}$$

We have:

$$\frac{n*}{n} - 1 = \frac{\sum_{i=1}^{l_n} \xi_i \tau_i}{n} - 1 = \frac{l_n}{n}\left[\frac{\sum_{i=1}^{l_n} (\xi_i \tau_i + \xi_i \mathbb{E}^*(\tau) - \xi_i \mathbb{E}^*(\tau)) - n}{l_n}\right]$$

$$= \frac{l_n}{n}\left[\frac{\sum_{i=1}^{l_n} \xi_i(\tau_i - \mathbb{E}^*(\tau))}{l_n}\right] + \frac{l_n}{n}\left[\frac{\sum_{i=1}^{l_n} \xi_i \mathbb{E}^*(\tau)}{l_n} - \frac{n}{l_n}\right]$$

$$= I + II. \tag{24}$$

We denote by $\mathbb{E}^*$ the expectation conditionally on $X_1, \ldots, X_n$. By the fact that τ_i are i.i.d. and using Chebyshev's inequality, we have:

$$
\begin{aligned}
\mathbb{P}_{\xi|X}(|I| > \epsilon) &\leq \epsilon^{-2} \left(\frac{l_n}{n}\right)^2 \frac{1}{l_n} \mathbb{E}_{\xi|X}(\xi_{1,l_n}(\tau_1 - \mathbb{E}^*(\tau)))^2 \\
&\leq 2\epsilon^{-2} \left(\frac{l_n}{n}\right) \frac{1}{n} \mathbb{E}(\xi_{1,l_n}^2) \frac{1}{l_n} \sum_{i=1}^{l_n} \tau_i^2 \to 0 \text{ in probability.}
\end{aligned}
$$

The last inequality follows using i), which implies that $\frac{l_n}{n} \to p$ and iii) where

$$
\frac{1}{l_n} \sum_{i=1}^{l_n} \tau_i^2 \to \mathbb{E}(\tau^2),
$$

for $\mathbb{E}(\xi_{1,l_n}^2) < \infty$. For II we have:

$$
\begin{aligned}
II &= \frac{l_n}{n} \left[\frac{\sum_{i=1}^{l_n} \xi_i \mathbb{E}^*(\tau)}{l_n} - \frac{n}{l_n} \right] = \frac{l_n}{n} \left[\mathbb{E}^*(\tau) - \frac{n}{l_n} \right] \text{ by (A1)} \\
&= \frac{l_n}{n} \left[\frac{1}{l_n} \sum_{i=1}^{l_n} (\tau_i - \mathbb{E}(\tau) + \mathbb{E}(\tau)) - \frac{n}{l_n} \right] \\
&= \frac{l_n}{n} \left[\frac{1}{l_n} \sum_{i=1}^{l_n} (\tau_i - \mathbb{E}(\tau)) \right] + \frac{l_n}{n} \left[\mathbb{E}(\tau) - \frac{n}{l_n} \right].
\end{aligned}
$$

The last equality converges to zero by the fact that $n/l_n \to \alpha = \mathbb{E}(\tau)$ and by iii)

$$
\frac{1}{l_n} \sum_{i=1}^{l_n} (\tau_i - \mathbb{E}(\tau)) \to 0.
$$

This proves Lemma 1. □

Proof of Theorem 3. For the weak convergence, we need to show the finite-dimensional convergence and the asymptotic equicontinuity. According to Proposition 2 and [6], the finite-dimensional convergence is considered if, for every fixed finite collection of functions $\{f_1, \ldots, f_k\} \subset \mathscr{F}$,

$$
\left(\binom{n}{m_1}^{-1/2} R_{l_n}^*(f_1), \ldots \binom{n}{m_k}^{-1/2} R_{l_n}^*(f_k) \right) \to \left(\mathbb{K}_{\mathbb{P}}(f_1), \ldots, \mathbb{K}_{\mathbb{P}}(f_k) \right),
$$

where $\mathbb{K}_P$ is the Gaussian chaos process. According to Cramér–Wold and the countability of $\mathscr{F}$, we only need to show that for any $f \in L_2^{c,m}(\mathbb{P})$,

$$
\sup_{\psi \in \mathrm{BL}} \left| \mathbb{E}\left[\psi\left(\binom{n}{m}^{-1/2} R_{l_n}^*(f) \right) \bigg| \{\mathbf{B}_i\} \right] - \mathbb{E}\psi(c \cdot \mathbb{K}_{\mathbb{P}}(f)) \right| \to 0 \text{ a.s.} \tag{25}
$$

By ([6] Section 4.2) and ([29] Section 2A), any $f \in L_2^{c,m}(\mathbb{P})$ can be expanded in $L_2(\mathbb{P}^m)$ by $f = \sum_{q=1}^{\infty} c_q h_m^{\psi_q}$, where $\{c_q\}$ is a sequence of real numbers and

$$
h_m^{\psi_q}(x_1, \ldots, x_m) \equiv \psi_q(x_1) \cdots \psi_q(x_m)
$$

for some bounded $\psi_q \in L_2^{c,1}(\mathbb{P})$. Fix $\epsilon > 0$. Then, there exists $Q_\epsilon \in \mathbb{N}$ such that with $f_n^\epsilon \equiv \sum_{q=1}^{Q_\epsilon} c_q h_m^{\psi_q}$,

$$
\|f - f_\epsilon\|_{L_2(P^m)} \leq \epsilon.
$$

The left-hand side of (25) can be further bounded by

$$
\sup_{\psi \in \mathrm{BL}} \left| \mathbb{E}\left[\psi\left(\binom{n}{m}^{-1/2} R_{l_n}^*(f) \right) \Big| \{\mathbf{B}_i\} \right] - \mathbb{E}\psi(c \cdot \mathbb{K}_{\mathbb{P}}(f)) \right|
$$

$$
\leq \sup_{\psi \in \mathrm{BL}} \left| \mathbb{E}\left[\psi\left(\binom{n}{m}^{-1/2} R_{l_n}^*(f) \right) \Big| \{\mathbf{B}_i\} \right] - \mathbb{E}\left[\psi\left(\binom{n}{m}^{-1/2} R_{l_n}^*(f^\epsilon) \right) \Big| \{\mathbf{B}_i\} \right] \right|
$$

$$
+ \sup_{\psi \in \mathrm{BL}} \left| \mathbb{E}\left[\psi\left(\binom{n}{m}^{-1/2} R_{l_n}^*(f^\epsilon) \right) \Big| \{\mathbf{B}_i\} \right] - \mathbb{E}\psi(c \cdot \mathbb{K}_{\mathbb{P}}(f^\epsilon)) \right|
$$

$$
+ \sup_{\psi \in \mathrm{BL}} \left| \mathbb{E}\psi(c \cdot \mathbb{K}_{\mathbb{P}}(f^\epsilon)) - \mathbb{E}\psi(c \cdot \mathbb{K}_{\mathbb{P}}(f)) \right|
$$

$$
\equiv (I) + (II) + (III). \tag{26}
$$

Let $\bar{f}^\epsilon \equiv f - f^\epsilon$; noting that ψ is bounded by one and using Lemma 1, we can replace l_n by $\varphi(n) = \left\lfloor \frac{n}{\mathbb{E}_A(\tau)} \right\rfloor$ which is deterministic. In the following, we denote by π a random permutation uniformly distributed over $\Sigma(n)$, the set of all permutations over $1, \ldots, n$. We have

$$
(I)^2 \leq \mathbb{E}^* \left| 2 \wedge \binom{n}{m}^{1/2} R_{l_n}^*(\bar{f}^\epsilon) \right|^2
$$

$$
\lesssim \mathbb{E}_{\xi|X}\mathbb{E}_R \left(1 \wedge n^{-m/2} \sum_{1 \leq i_1 \neq \ldots \neq i_m \leq \varphi(n)} \left(\xi_{\pi_{i_1}} - 1 \right) \cdots \left(\xi_{\pi_{i_m}} - 1 \right) \bar{f}_\varepsilon\left(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m} \right) \right)^2
$$

$$
\lesssim \sum_{\alpha_i \in \{1,2\}: \sum_{i=1}^l \alpha_i = 2m, \alpha_1 \geq \ldots \geq \alpha_l, 1 \leq l \leq m} \mathbb{E}_\xi^* \left[1 \wedge n^{-m/2} \mathbb{E}_R \left[\prod_{i=1}^l \left(\xi_{\pi_i} - 1 \right)^{\alpha_i} \right] \right.
$$

$$
\times \sum_{\substack{i_1 \neq \ldots \neq i_m, \\ i_1' \neq \ldots \neq i_m', \\ i_j = i_j', 1 \leq j \leq \max\{j: \alpha_j = 2\}}} \bar{f}^\epsilon\left(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_{m(n)}} \right) \bar{f}_n^\epsilon\left(\mathbf{B}_{i_1'}, \ldots, \mathbf{B}_{i_m'(n)} \right) \right]
$$

$$
\lesssim \sum_{\substack{\alpha_i \in \{1,2\}: \sum_{i=1}^l \alpha_i = 2m, \\ \alpha_1 \geq \ldots \geq \alpha_l, 1 \leq \ell \leq m}} \mathbb{E}\left[1 \wedge \frac{1}{n} \sum_{i=1}^n \left(\xi_i - 1 \right)^2 \right]^m
$$

$$
\times n^{-\ell} \sum_{\substack{i_1 \neq \ldots \neq i_m, \\ i_1' \neq \ldots \neq i_m', \\ i_j = i_j', 1 \leq j \leq \max\{j: \alpha_j = 2\}}} \bar{f}^\epsilon\left(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m} \right) \bar{f}^\epsilon\left(\mathbf{B}_{i_1'}, \ldots, \mathbf{B}_{i_m'} \right).
$$

We have, according to [43], for $(\xi_1, \ldots, \xi_n)$ a non-negative sequence of variables such that $\sum_{i=1}^n \xi_i = n$ and for $\pi = (\pi_1, \ldots, \pi_n)$ a random permutation of $\{1, \ldots, n\}$, for any $\ell \in \mathbb{N}$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{N}^\ell$,

$$
\left| \mathbb{E}_\pi \left[\prod_{i=1}^\ell \left(\xi_{\pi_i} - 1 \right)^{\alpha_i} \right] \right| \leq C_{l,\boldsymbol{\alpha}} n^{-\ell} \left[\sum_{i=1}^{l_n} \left(\xi_i - 1 \right)^2 \right]^{\sum_i \alpha_i/2}.
$$

Furthermore, according to [70,71], we have:

$$
n^{-\ell} \sum_{\substack{i_1 \neq \ldots \neq i_m, \\ i_1' \neq \ldots \neq i_m', \\ i_j = i_j', 1 \leq j \leq \max\{j: \alpha_j = 2\}}} \bar{f}^\epsilon\left(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m} \right) \bar{f}^\epsilon\left(\mathbf{B}_{i_1'}, \ldots, \mathbf{B}_{i_m'} \right)
$$

$$
\to_{\mathrm{a.s.}} \mathbb{E}_A(\tau)^{-\ell} \mathbb{E}\bar{f}^\epsilon\left(\mathbf{B}_1, \ldots, \mathbf{B}_m \right) \bar{f}^\epsilon\left(\mathbf{B}_1', \ldots, \mathbf{B}_m' \right)
$$

(where $\mathbf{B}_j = \mathbf{B}'_j$ for $1 \le j \le \max\{j : \alpha_j = 2\}$ and for $l_n/n \to \mathbb{E}_A(\tau)^{-1}$)

$$\le \mathbb{E}_A(\tau)^{-\ell} \mathbb{P}^m \bar{f}^{\epsilon^2} \le \epsilon^2. \quad \text{(Under Conditions (C.1) and (C.3).)}$$

Hence we have

$$\limsup_{n \to \infty}(I) \lesssim_{m,\xi} \epsilon, \text{ a.s.} \tag{27}$$

Now for the second term, we have:

$$\binom{n}{m}^{-1/2} R^*_{l_n}(f^\epsilon)$$

$$= \frac{1}{\binom{n}{m}^{1/2}} \sum_{q=1}^{Q_\epsilon} c_q \sum_{1 \le i_1 < \dots < i_m \le \varphi(n)} (\xi_{\pi_{i_1}} - 1) \cdots (\xi_{\pi_{i_m}} - 1) \psi_q(\mathbf{B}_{i_1}) \cdots \psi_q(\mathbf{B}_{i_m})$$

$$= \frac{\varphi(n)^{m/2}}{\binom{n}{m}^{1/2}} \sum_{q=1}^{Q_\epsilon} c_q \Re_m \left(\frac{1}{\varphi(n)^{1/2}} \sum_{i=1}^{\varphi(n)} (\xi_{\pi_i} - 1)\psi_q(\mathbf{B}_i), \dots, \frac{1}{\varphi(n)^{m/2}} \sum_{i=1}^{\varphi(n)} (\xi_{\pi_i} - 1)^m \psi_q^m(\mathbf{B}_i) \right)$$

$$\equiv (1 + o(1))(m!)^{1/2} \mathbb{E}_A(\tau)^{-m/2} \sum_{q=1}^{Q_\epsilon} c_q \Re_m(A^{(1)}_{\varphi(n),q}, \dots, A^{(m)}_{\varphi(n),q}),$$

where $\Re_m$ is the polynomial of degree m (see [6], p. 175):

$$\sum_{1 \le i_1 < \dots < i_m \le \varphi(n)} t_{i_1} \cdots t_{i_m} = \Re_m \left(\sum_{i=1}^{\varphi(n)} t_i, \sum_{i=1}^{\varphi(n)} t_i^2, \dots, \sum_{i=1}^{\varphi(n)} t_i^m \right). \tag{28}$$

As we mentioned before, this polynomial follows from Newton's inequality and allows us to show a polynomial function as a sum of monomials. All we need now is to check each argument of this polynomial function. $\square$

For $\ell = 1$: We first recall the following lemma from [53].

Lemma 3 ([53]). *Let $(a_1, \dots, a_n)$ be a vector and $(\xi_1, \dots, \xi_n)$ be a vector of exchangeable random variables. Suppose that*

$$\bar{a}_n = \frac{1}{n} \sum_{i=1}^n a_i = 0, \quad \frac{1}{n} \sum_{i=1}^n a_i^2 \to \sigma^2, \quad \lim_{M \to \infty} \limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^n a_i^2 \mathbb{1}_{\{|a_i| > M\}} = 0,$$

and

$$\bar{\xi}_n = \frac{1}{n} \sum_{i=1}^n \xi_i = 0, \quad \frac{1}{n} \sum_{i=1}^n \xi_i^2 \to_{\mathbb{P}_\xi} \alpha^2, \quad \frac{1}{n} \max_{1 \le i \le n} \xi_i^2 \to_{\mathbb{P}_\xi} 0.$$

Then,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n a_i \xi_i \to \mathcal{N}(0, \sigma^2 \alpha^2).$$

Applying Lemma 3 with $a_i \equiv \psi_q(\mathbf{B}_i) - \mathbb{P}_n \psi_q$ and ξ_i replaced by $\xi_{R_i} - 1$, we can see that

$$A^{(1)}_{\varphi(n),q} \to c \cdot \mathbb{G}_{\mathbb{P}}(\psi_q), \text{ a.s.,}$$

where $\mathbb{G}_{\mathbb{P}}$ is a Gaussian process defined on $L_2^{c,1}(\mathbb{P})$ with covariance

$$\mathbb{E}\mathbb{G}_{\mathbb{P}}(f)\mathbb{G}_{\mathbb{P}}(g) = \mathbb{P}(fg), \text{ for } f, g \in L_2^{c,1}(\mathbb{P}).$$

For $\ell = 2$: Note that

$$\mathbb{E}_\pi^{*,\xi}(A_{\varphi(n),q}^{(2)}) = \frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} (\xi_i - 1)^2 \cdot \frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} \psi_q^2(\mathbf{B}_i)$$

$$\to_{\mathbb{P}_{\nu,\xi}} c^2 \mathbb{E}\psi_q^2(\mathbf{B}_1) = c^2 \mathbb{E}_A\left[\left(\sum_{i=\mathscr{T}_0+1}^{\mathscr{T}_1} h_1(X_i)\right)^2\right], \text{ a.s.}$$

Furthermore,

$$\text{Var}^{*,\xi}(A_{\varphi(n),q}^{(2)})$$

$$= \mathbb{E}_\pi^{*,\xi}(A_{\varphi(n),q}^{(2)})^2 - \left(\mathbb{E}_\pi^{*,\xi}(A_{\varphi(n),q}^{(2)})\right)^2$$

$$= \mathbb{E}_\pi^{*,\xi}\left[\frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} (\xi_i - 1)^2 \psi_q^2(\mathbf{B}_{\pi_i})\right]^2 - \left[\frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} (\xi_i - 1)^2 \mathbb{P}_n \psi_q^2\right]^2$$

$$= \frac{1}{\varphi(n)^2} \sum_{i,j} (\xi_i - 1)^2 (\xi_j - 1)^2 \left[\mathbb{E}_\pi^* \psi_q^2(\mathbf{B}_{\pi_i}) \psi_q^2(\mathbf{B}_{\pi_j}) - (\mathbb{P}_n \psi_q^2)^2\right]$$

$$= \frac{1}{\varphi(n)^2} \sum_i (\xi_i - 1)^4 \left[\mathbb{E}_\pi^* \psi_q^4(\mathbf{B}_{\pi_i}) - (\mathbb{P}_n \psi_q^2)^2\right]$$

$$+ \frac{1}{\varphi(n)^2} \sum_{i \neq j} (\xi_i - 1)^2 (\xi_j - 1)^2 \left[\mathbb{E}_\pi^* \psi_q^2(\mathbf{B}_{\pi_i}) \psi_q^2(\mathbf{B}_{\pi_j}) - (\mathbb{P}_n \psi_q^2)^2\right]$$

$$\leq \frac{1}{\varphi(n)^2} \sum_i (\xi_i - 1)^4 \cdot \mathbb{P}_n \psi_q^4 + \frac{1}{\varphi(n)^2} \left(\sum_i (\xi_i - 1)^2\right)^2 \cdot \frac{1}{\varphi(n) - 1} \mathbb{P}_n \psi_q^4$$

$$\leq C \frac{1}{\varphi(n)^2} \sum_{i=1}^n (\xi_i - 1)^4 \cdot \mathbb{P}_n \psi_q^4$$

$$\leq C\|\psi_q\|_\infty^4 \frac{\max_i (\xi_i - 1)^2}{\varphi(n)} \cdot \frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} (\xi_i - 1)^2 \to_{\mathbb{P}_\xi} 0, \text{ a.s.}$$

The first inequality in the above display follows, since

$$\mathbb{E}_\pi^* \psi_q^2(\mathbf{B}_{\pi_i} \mathbf{B}_{\pi_i}) \psi_q^2(\mathbf{B}_{\pi_j}) - (\mathbb{P}_n \psi_q^2)^2$$

$$= \frac{1}{\varphi(n)(\varphi(n) - 1)} \left[\sum_{i \neq j} \psi_q^2(\mathbf{B}_{\pi_i}) \psi_q^2(\mathbf{B}_{\pi_j})\right] - (\mathbb{P}_n \psi_q^2)^2$$

$$\leq \frac{1}{\varphi(n) - 1} (\mathbb{P}_n \psi_q^2)^2 \leq \frac{1}{\varphi(n) - 1} \mathbb{P}_n \psi_q^4.$$

This shows that

$$A_{\varphi(n),q}^{(2)} \to_{\mathbb{P}_\xi} c^2 \mathbb{E}\psi_q^2 \text{ a.s.}$$

For $\ell \geq 3$:

$$\mathbb{E}_\pi^{*,\xi}|A_{\varphi(n),q}^{(\ell)}| \leq \frac{1}{\varphi(n)^{\ell/2}} \sum_{i=1}^{\varphi(n)} |\xi_i - 1|^\ell \cdot \frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} |\psi_q(\mathbf{B}_i)|^\ell$$

$$\leq \left(\frac{\max_i |\xi_i - 1|^2}{\varphi(n)}\right)^{\frac{\ell-2}{2}} \cdot \frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} |\xi_i - 1|^2 \cdot \|\psi_q\|_\infty$$

$$\to_{\mathbb{P}_\xi} 0, \text{ a.s.}$$

This shows that

$$A_{\varphi(n),q}^{(\ell)} \to_{\mathbb{P}_\xi} 0, \text{ a.s.}$$

Then, we have

$$R_m(A^{(1)}_{\varphi(n),q}, \ldots, A^{(m)}_{\varphi(n),q}) \to R_m(\mathbb{G}_{\mathbb{P}}(c\psi_q), \mathbb{E}(c\psi_q)^2, 0, \ldots, 0) = c(m!)^{-1/2} \cdot \mathbb{K}_{\mathbb{P}}(\psi_q) \text{ a.s.,}$$

where $\mathbb{K}_{\mathbb{P}}$ is the Gaussian chaos process defined on ($\oplus$ is the orthogonal sum in $L_2(E^\infty, \mathscr{E}^\infty, P^\infty)$)

$$\mathbb{R} \oplus L_2^{c,\mathbb{N}}(\mathbb{P}) \equiv \mathbb{R} \oplus \left(\oplus_{m=1}^\infty L_2^{c,m}(\mathbb{P}) \right).$$

Hence, it follows that, by linearity of $\mathbb{K}_{\mathbb{P}}$,

$$\binom{n}{m}^{-1/2} R^*_{l_n}(\bar{f}^\epsilon) \to c \cdot \mathbb{K}_{\mathbb{P}}(f^\epsilon), \text{ a.s.}$$

The last term in (26) follows from the definition of $\mathbb{K}_{\mathbb{P}}$

$$(III) \leq c\sqrt{\mathbb{E}\mathbb{K}_{\mathbb{P}}^2(\bar{f}^\epsilon)} \to 0 \quad (\epsilon \to 0). \tag{29}$$

All these final results give the finite-dimensional convergence.

Now, we take a step-by-step approach to establish stochastic equicontinuity. We assume that the class of functions must be bounded, so we suppose that $h \leq H$, for H an envelope. Throughout the following, we denote by

$$\mathscr{F}_\delta := \{f, g \in \mathscr{F} : d(f, g) \leq \delta\}.$$

Step 1

Let

$$Z_n^* := \binom{n^*}{m}^{1/2} [U_n^*(h) - \mathbb{E}^*(U_n^*(h))], \tag{30}$$

and

$$\check{T}^*_{l_n} := \binom{n}{m}^{-1/2} \binom{l_n}{m} \left[R^*_{l_n} - \mathbb{E}^*(R^*_{l_n}) \right]. \tag{31}$$

In this step, we must prove that the stochastic equicontinuity of the U-process implies that of the regenerative U-process. This is a consequence of 1, and for the weighted bootstrap Proposition 2 and part ii) of Lemma 1.

Step 2

Define

$$\check{T}^*_{l_n} := \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in I^m_{l_n}} (\xi_{i_1,l_n} - 1)\ldots(\xi_{i_m,l_n} - 1)\omega_h(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m})$$

and

$$\widetilde{T}^*_{l_n} = \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in I^m_{\mathbb{E}(l_n)}} (\xi_{i_1,l_n} - 1)\ldots(\xi_{i_m,l_n} - 1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m}).$$

Hypothesis: The stochastic equicontinuity of $\check{T}^*_{l_n}$ implies the stochastic equicontinuity of $\widetilde{T}^*_{l_n}$.

Proof. In order to prove the previous implication, we only need to show that:

$$\mathbb{P}^* \left(\left\| \widetilde{T}^*_{l_n} - \check{T}^*_{l_n} \right\|_{\mathscr{F}_\delta} > \epsilon \right)$$

$$\leq \quad \mathbb{P}^*\left(\left\|\binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in I^m_{\mathbb{E}(l_n)}} (\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_h(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > \epsilon\right).$$

Suppose that $l_n \leq \mathbb{E}(l_n)$, the opposite case can be treated in a similar way. We have

$$\mathbb{P}^*\left(\left\|\widetilde{T}^*_{l_n} - \check{T}^*_{l_n}\right\|_{\mathscr{F}_\delta} > \epsilon\right)$$

$$= \quad \mathbb{P}^*\left(\left\|\binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in I^m_{\mathbb{E}(l_n)}} (\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_h(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right.\right.$$

$$\left.\left. - \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in I^m_{l_n}} (\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_h(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > \epsilon\right)$$

Define $\mathfrak{I} := \big\{(i_1,\ldots,i_m) : 1 \leq i_1 < \ldots < i_m \leq \mathbb{E}(l_n) : i_j \neq i_k \text{ for } j \neq k,$
$$\text{such that } \exists \ell \in \{1,\ldots,m\} : l_n \leq i_\ell \leq \mathbb{E}(l_n)\big\}$$

$$\leq \quad \mathbb{P}^*\left(\left\|\binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in\mathfrak{I}} (\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_h(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > \epsilon\right)$$

$$= \quad \mathbb{P}^*\left(\left\|\binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in\mathfrak{I}} (\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_h(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m}) > \epsilon\right.\right.$$

$$\left.\left. \cap (|\mathbb{E}(l_n)-l_n| \leq n/4)\right\|_{\mathscr{F}_\delta}\right) + \mathbb{P}^*(|\mathbb{E}(l_n)-l_n| > n/4).$$

However, $|\mathbb{E}(l_n) - l_n| = O_{\mathbb{P}}(\sqrt{n})$ by Lemma 1, part i). Then, the exists a constant $K > 0$, such that for every $\epsilon > 0$,

$$\mathbb{P}^*(|\mathbb{E}(l_n)-l_n| > n/4) < \epsilon,$$

and the first expression in the previous expression is bounded by:

$$\mathbb{P}^*\left(\max_{M\leq n/2+\mathbb{E}(l_n)}\left\|\binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in\mathfrak{I}'} (\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > \epsilon\right)$$

where $\mathfrak{I}' := \{(i_1,\ldots,i_m) : 1 \leq i_1 < \ldots < i_m \leq \mathbb{E}(l_n), \exists \ell = 1,\ldots,m, \mathbb{E}(l_n) < i_\ell \leq M,$
$$i_j \neq i_k \text{ for } j \neq k\}$$

$$\leq \quad C_1\mathbb{P}^*\left(\left\|\binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m)\in\mathfrak{I}''} (\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > C_2\epsilon\right)$$

where $\mathfrak{I}''_m := \{(i_1,\ldots,i_m) : 1 \leq i_1 < \ldots < i_m \leq \mathbb{E}(l_n), \exists \ell = 1,\ldots,m,$
$$\mathbb{E}(l_n) < i_\ell \leq \mathbb{E}(l_n) + n/2, i_j \neq i_k \text{ for } j \neq k\}.$$

The last expression follows from the Montgomery–Smith inequality. Since

$$\mathbb{E}(l_n)/n \to \alpha^{-1},$$

the last expression matches the stochastic equicontinuity condition for $\widetilde{T}^*_{l_n}$. This proves this step. $\square$

Before passing to the next step, we introduce a new bootstrap sample. Define $\hat{\mathbf{B}}_i := \left\{ X_{\mathcal{T}_{i-1}+1}, \ldots, X_{\mathcal{T}_i} \right\}$ for $i = 1, \ldots, \mathbb{E}(l_n)$. Now, apply the weighted bootstrap procedure on the sample $\{\hat{\mathbf{B}}_i\}_{i=1}^{\mathbb{E}(l_n)}$. This new procedure is the same as the old one for $\mathbf{B}_i$, but we aim here to replace the random quantity l_n with a deterministic one, which is $\mathbb{E}(l_n)$.

Step 3

Define:

$$\widehat{T}^*_{l_n} = \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in I^m_{\mathbb{E}(l_n)}} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \omega_{\widetilde{h}}(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m})$$

Hypothesis: The stochastic equicontinuity of $\widetilde{T}^*_{l_n}$ implies the stochastic equicontinuity of $\widehat{T}^*_{l_n}$.

Proof. First case: $l_n \leq \mathbb{E}(l_n)$:

In this case, all of the terms in the following computation should be multiplied with $\mathbb{1}_{(l_n \leq \mathbb{E}(l_n))}$. We leave it out to keep the already complex notation simple. Define

$$
\begin{aligned}
\mathscr{A}_n &:= \{\mathbf{B}_1, \ldots, \mathbf{B}_{l_n}\} \\
\overrightarrow{T}^*_{l_n} &:= \widehat{T}^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n} + T^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n^c}.
\end{aligned}
$$

$\overrightarrow{T}^*_{l_n}$ is well defined, i.i.d., and has the same distribution as $T^*_{l_n}$ and $(i_1, \ldots, i_m) \in I^m_{\mathbb{E}(l_n)}$. Hence, if we show that:

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}^* \left(\left\| \overrightarrow{T}^*_{l_n} \right\|_{\mathscr{F}_\delta} > \epsilon \right) = 0 \quad \text{in probability,}$$

then the stochastic equicontinuity of $\widetilde{T}^*_{l_n}$ is established. However, we aim to approximate the one of $\widehat{T}^*_{l_n}$. In order to achieve our goal, it is sufficient to estimate:

$$
\begin{aligned}
& \left\| \overrightarrow{T}^*_{l_n} - \widehat{T}^*_{l_n} \right\|_{\mathscr{F}_\delta} \\
={} & \left\| \widehat{T}^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n} + T^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n^c} - \left[\widehat{T}^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n} + \widehat{T}^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n^c} \right] \right\|_{\mathscr{F}_\delta} \\
\leq{} & \left\| T^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n^c} \right\|_{\mathscr{F}_\delta} + \left\| \widehat{T}^*_{l_n} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n^c} \right\|_{\mathscr{F}_\delta} \\
:={} & I_n + II_n.
\end{aligned}
$$
(32)

For I_n: Let

$$S^*_n := \sum_{(i_1,\ldots,i_m) \in I^m_{\mathbb{E}(l_n)}} \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n^c},$$

conditioned on the sample, we have:

$$
\begin{aligned}
& \mathscr{L} \left(\sum_{(i_1,\ldots,i_m) \in I^m_{\mathbb{E}(l_n)}} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \omega_{\widetilde{h}}(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m}) \mathbb{1}_{(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m}) \in \mathscr{A}_n^c} \right) \\
={} & \mathscr{L} \left(\sum_{(i_1,\ldots,i_m) \in I^m_{S^*_n}} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \omega_{\widetilde{h}}(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m}) \right).
\end{aligned}
$$

Hence,

$$\mathbb{P}^*(I_n > \epsilon)$$

$$= \mathbb{P}^*\left(\binom{n}{m}^{-1/2}\left\|\sum_{(i_1,\ldots,i_m)\in I_{S_n^*}^m}(\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > \epsilon\right)$$

$$= \mathbb{P}^*\left(\left\|\binom{n}{m}^{-1/2}\sum_{(i_1,\ldots,i_m)\in I_{S_n^*}^m}(\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m}) > \epsilon\right.\right.$$

$$\cap\left(S_n^* \le K\sqrt{n})\Big\|_{\mathscr{F}_\delta}\right) + \mathbb{P}^*\left(S_n^* > K\sqrt{n}\right)$$

$$\le \mathbb{P}^*\left(\max_{M\le K\sqrt{n}}\left\|\binom{n}{m}^{-1/2}\sum_{(i_1,\ldots,i_m)\in\mathfrak{J}'}(\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > \epsilon\right)$$

$$+\mathbb{P}^*\left(S_n^* > K\sqrt{n}\right), \text{for any } K > 0,$$

where $\mathfrak{J}' := \{(i_1,\ldots,i_m) : 1 \le i_1 < \ldots < i_m \le S_n^*, \exists \ell = 1,\ldots,m, S_n^* < i_\ell \le M,$

$$i_j \ne i_k \text{ for } j \ne k\}.$$

$$\le C_1\mathbb{P}^*\left(\left\|\binom{n}{m}^{-1/2}\sum_{(i_1,\ldots,i_m)\in\mathfrak{J}_m''}(\xi_{i_1,l_n}-1)\ldots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > C_2\epsilon\right)$$

$$+\mathbb{P}^*\left(S_n^* > K\sqrt{n}\right), \text{for any } K > 0, \tag{33}$$

where

$$\mathfrak{J}_m'' := \{(i_1,\ldots,i_m) : 1 \le i_1 < \ldots < i_m \le S_n^*, \exists \ell = 1,\ldots,m,$$

$$S_n^* < i_\ell \le K\sqrt{n}, i_j \ne i_k \text{ for } j \ne k\}.$$

For *n* large enough, we need to show that there exists $K > 0$ such that

$$\mathbb{P}^*\left(S_n^* > K\sqrt{n}\right) \to 0.$$

As $\mathbb{1}_{\hat{\mathbf{B}}_i\in\mathscr{A}_n^c}$ are i.i.d and bounded,

$$\frac{S_n^* - \mathbb{E}(S_n^*)}{\sqrt{\mathbb{E}(l_n)}} \to N(0,\eta^2) \text{ in probability.}$$

therefore, we can find $\mathscr{M} > 0$ such that

$$\mathbb{P}^*\left(S_n^* > \mathbb{E}(S_n^*) + \mathscr{M}\sqrt{n}\right) < \epsilon.$$

However,

$$\mathbb{E}(S_n^*) = \mathbb{E}(l_n)\mathbb{P}^*(\hat{\mathbf{B}}_i^* \in \mathscr{A}_n^c) = \mathbb{E}(l_n) - l_n = O_\mathbb{P}(\sqrt{n}),$$

by Lemma 1 i), then

$$\mathbb{P}^*\left(S_n^* > K\sqrt{n}\right) \to 0.$$

Then, we only need to estimate the first part in (33). Define the following bootstrap procedure: let $\overrightarrow{\mathbf{B}}_i := \left\{X_{\mathscr{T}_{i-1}+1},\ldots,X_{\mathscr{T}_i}, 0, 0,\ldots\right\}$ and let $\overrightarrow{\mathscr{F}}$ be a class of function, related to the class of functions $\mathscr{F}$, such that, for every $\overrightarrow{\omega}_h \in \overrightarrow{\mathscr{F}}$:

$$
\begin{cases}
\overrightarrow{\omega}_h(\overrightarrow{\mathbf{B}}_1, \overrightarrow{\mathbf{B}}_2, \cdots, \overrightarrow{\mathbf{B}}_k) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} h(x_{i_1}, \ldots, x_{i_k}) \mathbb{1}_{x_k \neq 0} & \text{if defined,} \\
\infty & \text{otherwise.}
\end{cases}
\tag{34}
$$

It is classical that $\{\overrightarrow{B}_i\}$ are i.i.d., applying the same bootstrap method of Algorithm 1. This new sample allows us to enlarge and bound (33) by

$$
\mathbb{P}^* \left(\sup_{\overrightarrow{h} \in \overrightarrow{\mathscr{H}}} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}_m''} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \overrightarrow{\omega}_{\overrightarrow{h}}(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right| > \epsilon \right),
\tag{35}
$$

where $\overrightarrow{h} \in \overrightarrow{\mathscr{H}} = \{\overrightarrow{\omega}_{\overrightarrow{f}} - \overrightarrow{\omega}_{\overrightarrow{g}}, \overrightarrow{f}, \overrightarrow{g} \in \overrightarrow{\mathscr{F}}\}$ and the corresponding class $\mathscr{H} = \{\omega_f - \omega_g, f, g \in \mathscr{F}\}$, with envelope $\widetilde{F}$ and F, respectively. To estimate the last expression, we use *bracketing*. Define the *bracket* $[f^\ell, f^u]$ by:

$$
[f^\ell, f^u] := \{f \in \mathscr{F} : f^\ell \leqslant f \leqslant f^u\},
$$

and the *bracketing entropy number* by $N_1(\gamma, \mathscr{F}, \mathbb{P})$, which denotes the minimal number $N \geq 1$ for which there exist functions $f_1^\ell, \cdots, f_N^\ell$ and $f_1^u, \cdots, f_N^u$ such that:

$$
\begin{cases}
\mathscr{F} \subset \bigcup_{k=1}^{N} \left[f_k^\ell, f_k^u \right], \\
\int_S \left(f_k^u - f_k^\ell \right) \mathbb{P} \leq \gamma.
\end{cases}
\tag{36}
$$

For the class of functions $\mathscr{H}$, consider the bracket $[h^\ell, h^u]$, such that $\mathbb{E}^*(h^\ell, h^u) \leqslant \gamma$, where $\gamma > 0$ and it is determined later. In this framework, the bracketing entropy number is $N^*(\gamma) := N_1(\gamma, \mathscr{H}, \mathbb{D}^*)$, for

$$
\mathbb{D}^* = \binom{l_n}{m}^{-1} \sum_{(i_1,\ldots,i_m) \in I_{l_n}^m} \xi_{i_1,l_n} \ldots \xi_{i_m,l_n} \, \delta_{(\overrightarrow{\mathbf{B}}_{i_1},\ldots,\overrightarrow{\mathbf{B}}_{i_m})}.
$$

Hence, we have the following inequalities

$$
\sup_{\overrightarrow{h} \in \overrightarrow{\mathscr{H}}} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}_m''} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \overrightarrow{\omega}_{\overrightarrow{h}}(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right|
$$

$$
\leq \max_{k \leq N^*(\gamma)} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}_m''} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \left(h_k^u - h_k^\ell \right)(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right|
$$

$$
\leq \max_{k \leq N^*(\gamma)} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}_m''} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right.
$$

$$
\left. \left[h_k^u(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) - \mathbb{E}^* \left(h_k^u(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right|
$$

$$
+ \max_{k \leq N^*(\gamma)} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}_m''} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right.
$$

$$
\left. \left[h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) - \mathbb{E}^* \left(h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right|
$$

$$+ \max_{k \leq N^*(\gamma)} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right.$$

$$\left. \left[\mathbb{E}^* \left(h_k^u(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) - \mathbb{E}^* \left(h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right|$$

$$:= \quad I_A + I_B + I_C. \tag{37}$$

Treating each term, keeping in mind Condition (A.1), i.e., $\sum_{i=1}^n \xi_i = n$, we have

$$I_C := \max_{1 \leq k \leq N^*(\gamma)} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right.$$

$$\left. \left[\mathbb{E}^* \left(h_k^u(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) - \mathbb{E}^* \left(h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right|$$

$$\leq \quad \gamma \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right|$$

$$= \quad \gamma \left| \binom{n}{m}^{-1} \sum_{\mathbf{i}_{\{1,\ldots,m-1\}}} \sum_{j=1}^m \prod_{k=1,k\neq j}^m \xi_{i_k,n} \left\{ \sum_{i_j=1}^n (\xi_{i_j,n} - 1) \right\} \right| = 0,$$

and

$$\mathbb{P}(I_B > \epsilon) := \mathbb{P}\left(\max_{k \leq N^*(\gamma)} \left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right. \right.$$

$$\left. \left. \left[h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) - \mathbb{E}^* \left(h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right| > \epsilon \right)$$

$$\leq \quad N^*(\gamma) \max_{k \leq N^*(\gamma)} \mathbb{P}\left(\left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right. \right.$$

$$\left. \left. \left[h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) - \mathbb{E}^* \left(h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right| > \epsilon \right)$$

$$(\text{For } h_k^\ell = h_k^\ell \mathbb{1}_{h_k^\ell \leq M_n} + h_k^\ell \mathbb{1}_{h_k^\ell > M_n})$$

$$\leq \quad N^*(\gamma) \max_{k \leq N^*(\gamma)} \mathbb{P}\left(\left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right. \right.$$

$$\left. \left. \left[h_k^\ell \mathbb{1}_{h_k^\ell \leq M_n}(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) - \mathbb{E}^* \left(h_k^\ell \mathbb{1}_{h_k^\ell \leq M_n}(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right| > \epsilon \right)$$

$$+ N^*(\gamma) \max_{k \leq N^*(\gamma)} \mathbb{P}\left(\left| \binom{n}{m}^{-1/2} \sum_{(i_1,\ldots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \ldots (\xi_{i_m,l_n} - 1) \right. \right.$$

$$\left. \left. \left[h_k^\ell \mathbb{1}_{h_k^\ell > M_n}(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) - \mathbb{E}^* \left(h_k^\ell \mathbb{1}_{h_k^\ell > M_n}(\overrightarrow{\mathbf{B}}_{i_1}, \ldots, \overrightarrow{\mathbf{B}}_{i_m}) \right) \right] \right| > \epsilon \right)$$

$$\leq \quad \frac{N^*(\gamma)}{\epsilon^2 \binom{n}{m}^{1/2}} \mathbb{E}^* \left((\xi_{1,l_n} - 1) \ldots (\xi_{m,l_n} - 1) \left[h_k^\ell \mathbb{1}_{h_k^\ell \leq M_n} - \mathbb{E}^* \left(h_k^\ell \mathbb{1}_{h_k^\ell \leq M_n} \right) \right] \right)^2$$

$$+ \frac{N^*(\gamma)}{\epsilon} \mathbb{E}^* \left| (\xi_{1,l_n} - 1) \ldots (\xi_{m,l_n} - 1) \left[h_k^\ell \mathbb{1}_{h_k^\ell > M_n} - \mathbb{E}^* \left(h_k^\ell \mathbb{1}_{h_k^\ell > M_n} \right) \right] \right|$$

$$
\begin{aligned}
\leq \quad & \frac{N^*(\gamma)}{\epsilon^2\binom{n}{m}^{1/2}} \mathbb{E}_\xi\left(\prod_{i=1}^m (\xi_{i,l_n}-1)^2\right) \mathbb{E}^*\left(\left[h_k^\ell \mathbb{1}_{h_k^\ell \leq M_n} - \mathbb{E}^*\left(h_k^\ell \mathbb{1}_{h_k^\ell \leq M_n}\right)\right]\right)^2 \\
& + \frac{N^*(\gamma)}{\epsilon}\mathbb{E}^*\left|(\xi_{1,l_n}-1)\dots(\xi_{m,l_n}-1)\left[h_k^\ell \mathbb{1}_{h_k^\ell > M_n} - \mathbb{E}^*\left(h_k^\ell \mathbb{1}_{h_k^\ell > M_n}\right)\right]\right| \\
\leq \quad & \frac{N^*(\gamma)}{\epsilon^2\binom{n}{m}^{1/2}}c^2 \times 4M_n^2 + \frac{N^*(\gamma)}{\epsilon}\mathbb{E}^*\left|(\xi_{1,l_n}-1)\dots(\xi_{m,l_n}-1)h_k^\ell \mathbb{1}_{h_k^\ell > M_n}\right| \\
\leq \quad & \frac{N^*(\gamma)}{\epsilon^2\binom{n}{m}^{1/2}}c^2 \times 4M_n^2 + \frac{2N^*(\gamma)}{\epsilon}\mathbb{E}^*\left|(\xi_{1,l_n}-1)\dots(\xi_{m,l_n}-1)\widetilde{F}\mathbb{1}_{\widetilde{F} > M_n}\right| \\
\leq \quad & \frac{N^*(\gamma)}{\epsilon^2\binom{n}{m}^{1/2}}c^2 \times 4M_n^2 \\
& + \frac{4N^*(\gamma)}{\epsilon}\binom{l_n}{m}^{-1}\sum_{(i_1,\dots,i_m)\in I_{l_n}^m}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\left|\widetilde{F}\mathbb{1}_{\widetilde{F} > M_n}(\overrightarrow{\mathbf{B}}_{i_1},\dots,\overrightarrow{\mathbf{B}}_{i_m})\right|,
\end{aligned}
\tag{38}
$$

yet, $\overrightarrow{\mathbf{B}}_i$ are i.i.d. and and $\mathbb{E}\left(\widetilde{F}\right) = \mathbb{E}(\tau)^m\mathbb{E}F < \infty$, so for any $M_n \nearrow \infty$, we have

$$
\binom{\mathbb{E}(l_n)}{m}^{-1}\sum_{(i_1,\dots,i_m)\in I_{\mathbb{E}(l_n)}^m}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\widetilde{F}\mathbb{1}_{\widetilde{F}>M_n}(\overrightarrow{\mathbf{B}}_{i_1},\dots,\overrightarrow{\mathbf{B}}_{i_m}) \to 0 \;\; a.s.
$$

Using the same argument as in part iii) of Lemma 1, we can prove that

$$
\binom{l_n}{m}^{-1}\sum_{(i_1,\dots,i_m)\in I_{l_n}^m}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\left|\overrightarrow{F}\mathbb{1}_{\overrightarrow{F}>M_n}(\overrightarrow{\mathbf{B}}_{i_1},\dots,\overrightarrow{\mathbf{B}}_{i_m})\right| \to 0 \;\; in \;\; probability.
$$

Then, it remains to find that, for every fixed $\gamma > 0$, $N^*(\gamma)$ is bounded in probability, as the last expression in (38) does not depend on k. It is interesting to note that $N_1(\gamma, \overrightarrow{\mathscr{H}}, \mathbb{P})$ is finite, due to the boundness of $\overrightarrow{\mathscr{H}}$ by $2F$ with $\mathbb{E}\,\overrightarrow{F}\,(\overrightarrow{\mathbf{B}}) < \infty$ and the fact that $\overrightarrow{\mathbf{B}}_i$ are i.i.d. and discrete random variables. Under the norm $L_1(\mathbb{P})$, define $\gamma/2$ brackets $h_1^\ell, \cdots h_{N(\gamma/2)}^\ell$ and $h_1^u, \cdots, h_{N(\gamma/2)}^u$. Observe that

$$
\max_{j \leq N(\gamma/2)}\left|\binom{l_n}{m}^{-1/2}\sum_{(i_1,\dots,i_m)\in I_{l_n}^m}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\left(h_j^u - h_j^\ell\right)(\overrightarrow{\mathbf{B}}_{i_1},\dots,\overrightarrow{\mathbf{B}}_{i_m})\right|, \tag{39}
$$

converges to zero in probability, and $N(\gamma/2)$ does not depend on n. That implies that $N^*(\gamma) \leq N(\gamma/2) < \infty$ in probability. Replacing h^ℓ by h^u, I_A is identical to I_B, i.e., I_A also converges to zero in probability. This proves the convergence of I_n to zero in probability.

For II_n**:** In the same manner, let

$$
S_n^* := \sum_{(i_1,\dots,i_m)\in I_{\mathbb{E}(l_n)}^m}\mathbb{1}_{(\hat{\mathbf{B}}_{i_1},\dots,\hat{\mathbf{B}}_{i_m})\in\mathscr{A}_n^c}.
$$

Define a new bootstrap sample $\{\mathbf{B}_i^{**}\}$ in $i = l_n + 1, \dots, \mathbb{E}(l_n)$. Clearly, the new sample is well-defined since we assumed at the beginning that $l_n \leq \mathbb{E}(l_n)$, and it is defined independently from $\mathbf{B}_i^*$ and $\hat{\mathbf{B}}_i^*$. In this case:

$$
\mathscr{L}\left(\sum_{(i_1,\dots,i_m)\in I_{\mathbb{E}(l_n)}^m}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\hat{\mathbf{B}}_{i_1},\dots,\hat{\mathbf{B}}_{i_m})\mathbb{1}_{(\hat{\mathbf{B}}_{i_1},\dots,\hat{\mathbf{B}}_{i_m})\in\mathscr{A}_n^c}\right)
$$

$$= \mathscr{L}\left(\sum_{(i_1,\dots,i_m)\in I_{S_n^*}^m}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1}^{**},\dots,\mathbf{B}_{i_m}^{**})\right).$$

Hence, as in (33), we have:

$$\mathbb{P}^*(II_n > \epsilon)$$

$$= \mathbb{P}^{**}\left(\binom{n}{m}^{-1/2}\left\|\sum_{(i_1,\dots,i_m)\in I_{S_n^*}^m}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\dots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > \epsilon\right)$$

$$\leq C_1\mathbb{P}^{**}\left(\left\|\binom{n}{m}^{-1/2}\sum_{(i_1,\dots,i_m)\in\mathfrak{I}_m''}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\dots,\mathbf{B}_{i_m})\right\|_{\mathscr{F}_\delta} > C_2\epsilon\right)$$

$$+\mathbb{P}^*\left(S_n^* > K\sqrt{n}\right), \text{ for any } K > 0, \tag{40}$$

where

$$\mathfrak{I}_m'' := \{(i_1,\dots,i_m) : 1 \leq i_1 < \dots < i_m \leq S_n^*, \exists \ell = 1,\dots,m,$$
$$S_n^* < i_\ell \leq K\sqrt{n}, i_j \neq i_k \text{ for } j \neq k\}.$$

Using the same bootstrap procedure defined previously for I_n, let

$$\overrightarrow{\mathbf{B}}_i := \left\{X_{\mathscr{T}_{i-1}+1},\dots,X_{\mathscr{T}_i},0,0,\dots\right\},$$

for $i = l_n + 1,\dots,\mathbb{E}(l_n)$, and let $\overrightarrow{\mathscr{F}}$ be a class of function such that, for every $\overrightarrow{\omega}_h \in \overrightarrow{\mathscr{F}}$:

$$\begin{cases} \overrightarrow{\omega}_h(\overrightarrow{\mathbf{B}}_1,\overrightarrow{\mathbf{B}}_2,\cdots,\overrightarrow{\mathbf{B}}_k) = \sum_{i_1=1}^{\infty}\cdots\sum_{i_k=1}^{\infty}h(x_{i_1},\dots,x_{i_k})\mathbb{1}_{x_k\neq 0} & \text{if defined} \\ \infty & \text{otherwise.} \end{cases} \tag{41}$$

It is classical that $\{\overrightarrow{\mathbf{B}}_i\}$ are i.i.d., applying the same bootstrap method of Algorithm 1. This new sample allows us to enlarge and bound (33) by

$$\mathbb{P}^{**}\left(\sup_{\overrightarrow{h}\in\overrightarrow{\mathscr{H}}}\left|\binom{n}{m}^{-1/2}\sum_{(i_1,\dots,i_m)\in\mathfrak{I}_m''}(\xi_{i_1,l_n}-1)\dots(\xi_{i_m,l_n}-1)\overrightarrow{\omega}_{\overrightarrow{h}}(\overrightarrow{\mathbf{B}}_{i_1},\dots,\overrightarrow{\mathbf{B}}_{i_m})\right| > \epsilon\right), \tag{42}$$

where

$$\overrightarrow{h} \in \overrightarrow{\mathscr{H}} = \{\overrightarrow{\omega}_{\overrightarrow{f}} - \overrightarrow{\omega}_{\overrightarrow{g}}, \overrightarrow{f},\overrightarrow{g}\in\overrightarrow{\mathscr{F}}\}$$

corresponding to the class

$$\mathscr{H} = \{\omega_f - \omega_g, f,g\in\mathscr{F}\},$$

with envelope $\overrightarrow{F}$ and F, respectively. As before, for the class of functions $\mathscr{H}$, consider the bracket $[h^\ell, h^u]$, such that

$$\mathbb{E}^{**}(h^\ell, h^u) \leqslant \gamma,$$

where $\gamma > 0$ and it is determined later. In this framework, the bracketing entropy number is $N^{**}(\gamma) := N_1(\gamma,\mathscr{H},\mathbb{D}^{**})$, for

$$\mathbb{D}^{**} = \binom{\mathbb{E}(l_n)-l_n}{m}^{-1}\sum_{(i_1,\dots,i_m)\in I_{\mathbb{E}(l_n)-l_n}^m}\xi_{i_1,l_n}\dots\xi_{i_m,l_n}\,\delta_{(\overrightarrow{\mathbf{B}}_{i_1},\dots,\overrightarrow{\mathbf{B}}_{i_m})}.$$

Following the same arguments from Equations (37) through (38), we can find that (42) is

$$
\begin{aligned}
\leq \quad & N^{**}(\gamma) \max_{k \leq N^{**}(\gamma)} \mathbb{P}\Bigg(\Bigg| \binom{n}{m}^{-1/2} \sum_{(i_1,\dots,i_m) \in \mathfrak{I}''_m} (\xi_{i_1,l_n} - 1) \dots (\xi_{i_m,l_n} - 1) \\
& \Big[h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \dots, \overrightarrow{\mathbf{B}}_{i_m}) - \mathbb{E}^{**}\big(h_k^\ell(\overrightarrow{\mathbf{B}}_{i_1}, \dots, \overrightarrow{\mathbf{B}}_{i_m}) \big) \Big] \Bigg| > \epsilon \Bigg) \\
\leq \quad & \frac{N^{**}(\gamma)}{\epsilon^2 \binom{n}{m}^{1/2}} c^2 \times 4M_n^2 + \frac{2N^{**}(\gamma)}{\epsilon} \mathbb{E}^{**} \Big| (\xi_{1,l_n} - 1) \dots (\xi_{m,l_n} - 1) \widetilde{F} \mathbb{1}_{\widetilde{F} > M_n} \Big| \\
\leq \quad & \frac{N^{**}(\gamma)}{\epsilon^2 \binom{n}{m}^{1/2}} c^2 \times 4M_n^2 + \frac{4N^{**}(\gamma)}{\epsilon} \\
& \binom{\mathbb{E}(l_n) - l_n}{m}^{-1} \sum_{(i_1,\dots,i_m) \in I^m_{\mathbb{E}(l_n) - l_n}} (\xi_{i_1,l_n} - 1) \dots (\xi_{i_m,l_n} - 1) \Big| \widetilde{F} \mathbb{1}_{\widetilde{F} > M_n}(\overrightarrow{\mathbf{B}}_{i_1}, \dots, \overrightarrow{\mathbf{B}}_{i_m}) \Big|.
\end{aligned}
$$

(43)

Here, we must pay attention to the randomness of N^{**} which depends on n. According to Lemma 1 i), we can see that $|\mathbb{E}(l_n) - l_n| \to \infty$ in probability, under the assumption that $l_n < \mathbb{E}(l_n)$. Now, using the same treatment of I_n, and for $M_n := n^{1/3}$ (to provide the convergence of M_n to ∞), as in [47], this allows the convergence of (43) to zero in probability. Estimating now N^{**} by considering the same $\gamma/2$ brackets $h_1^\ell, \cdots h_{N(\gamma/2)}^\ell$ and $h_1^u, \cdots, h_{N(\gamma/2)}^u$, we have $N^{**}(\gamma) < N(\gamma/2)$, which does not depend on n. Then, II_n is proved. Following the same footsteps, we can prove the case where $l_n > \mathbb{E}(l_n)$. This proves Step 3. $\qquad \square$

The end of the previous step yields that we only need to show the stochastic equicontinuity of $\widehat{T}_{l_n}^*$, where the number of blocks is replaced by the deterministic quantity $\mathbb{E}(l_n)$. In order to achieve the equicontinuity of this statistic, Lemma 2 shows that it is sufficient to prove that:

$$
\mathbb{E} \Bigg\| \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m(n)} \epsilon_{i_1}^{(1)} \cdots \epsilon_{i_m}^{(m)} \omega_{\overrightarrow{h}}\Big(\hat{\mathbf{B}}_{i_1}^{(1)}, \dots, \hat{\mathbf{B}}_{i_m}^{(m)} \Big) \Bigg\|_{\overrightarrow{\mathscr{H}}}
$$
$$
\leq a(\ell_1, \dots, \ell_m) \Bigg(\prod_{k=1}^m \ell_k \Bigg)^{1/2}
$$

for all $1 \leq \ell_1, \dots, \ell_m \leq n$. We begin to define the distance:

$$
e_\ell^2(f, g) \equiv \frac{1}{\prod_{k=1}^m \ell_k} \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \omega_{\overrightarrow{h}}^2 \Big(\hat{\mathbf{B}}_{i_1}^{(1)}, \dots, \hat{\mathbf{B}}_{i_m}^{(m)} \Big),
$$

defined in L_2, associated with the Rademacher process

$$
\Bigg\{ \frac{1}{(\prod_{k=1}^m \ell_k)^{1/2}} \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \varepsilon_{i_1}^{(1)} \cdots \varepsilon_{i_m}^{(m)} \omega_{\overrightarrow{h}} \Big(\hat{\mathbf{B}}_{i_1}^{(1)}, \dots, \hat{\mathbf{B}}_{i_m}^{(m)} \Big) : \overrightarrow{h} \in \overrightarrow{\mathscr{H}} \mid \hat{\mathbf{B}}_1, \dots, \hat{\mathbf{B}}_m \Bigg\}.
$$

Take $\|f\|_\ell^2 \equiv e_\ell^2(f, 0)$ and

$$
r_\ell(\delta) \equiv \sup_{f \in \overrightarrow{\mathscr{F}}_\delta} \|f\|_\ell^2.
$$

Using Corollary A1, we have

$$\mathbb{E}\left\|\frac{1}{(\prod_{k=1}^m \ell_k)^{1/2}} \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \varepsilon_{i_1}^{(1)} \cdots \varepsilon_{i_m}^{(m)} \omega_{\overrightarrow{h}}\left(\hat{\mathbf{B}}_{i_1}^{(1)}, \ldots, \hat{\mathbf{B}}_{i_m}^{(m)}\right)\right\|_{\overrightarrow{\mathcal{H}}}$$

$$\leq C \int_0^{r_\ell(\delta)} (\log \mathcal{N}(\varepsilon, \mathcal{F}, e_\ell))^{m/2} \, d\varepsilon$$

$$= C\|F\|_\ell \cdot \int_0^{r_\ell(\delta)/\|F\|_\ell} (\log \mathcal{N}(\varepsilon\|F\|_\ell, \mathcal{F}, e_\ell))^{m/2} \, d\varepsilon$$

$$\leq C\|F\|_\ell \cdot \int_0^{r_\ell(\delta)/\|F\|_\ell} \left(\sup_Q \log \mathcal{N}\left(\varepsilon\|F\|_{L_2(Q)}, \mathcal{F}, L_2(Q)\right)\right)^{m/2} \, d\varepsilon. \tag{44}$$

Assuming that $F \geq 1$, the upper bound in the integral can be replaced by $r_\ell(\delta)$. The following proposition is necessary for the following.

Proposition 3 ([46]). *Let $\{X_i\}$ be i.i.d. random variables with law $\mathbb{P}$. Let $\mathcal{H}$ be a class of measurable real-valued functions defined on $(\mathcal{X}^m, \mathcal{A}^m)$ with an $\mathbb{P}^m$-integrable envelope such that the following holds: for any fixed $\delta > 0, M > 0, 1 \leq k \leq m$,*

$$\max_{1 \leq j' \leq k} \mathbb{E}\left(\frac{\log \mathcal{N}\left(\delta, (\pi_k \mathcal{H})_M, e_{\ell, j'}\right)}{\ell_{j'}}\right)^{1/2} \to 0 \tag{45}$$

holds for any $\ell_1 \wedge \cdots \wedge \ell_k \to \infty$. Here for $\ell = (\ell_1, \ldots, \ell_k)$ and $\{X_i\}_{i=1}^\infty$,

$$e_{\ell, j'}(f, g) \equiv \frac{1}{\ell_{j'}} \sum_{i_{j'}=1}^{\ell_{j'}} \left|\frac{1}{\prod_{j \neq j'} \ell_j} \sum_{1 \leq i_j \leq \ell_j : j \neq j'} (f - g)\left(X_{i_1}, \ldots, X_{i_k}\right)\right|$$

and

$$(\pi_k \mathcal{H})_M \equiv \{h \mathbb{1}_{H_k \leq M} : h \in \pi_k \mathcal{H}\},$$

where H_k is an envelope for $\pi_k \mathcal{H}$. Then,

$$\sup_{h \in \mathcal{H}} \left|\frac{1}{\prod_{k=1}^m \ell_k} \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \left(h\left(X_{i_1}, \ldots, X_{i_m}\right) - \mathbb{P}^m h\right)\right| \to 0$$

in L_1 as $\ell_1 \wedge \ldots \wedge \ell_m \to \infty$. The above equation can be replaced by the decoupled version.

By this proposition, $\|F\|_\ell \to_{\mathbb{P}} \|F\|_{L_2(P)}$ as $\ell_1 \wedge \ldots \wedge \ell_m \to \infty$, therefore, it suffices to get $r_\ell(\delta) \to_p 0$ as $\ell_1 \wedge \ldots \wedge \ell_m \to \infty$ and $\delta \to 0$. It is obvious that all that is left to do now is to demonstrate that

$$\sup_{f \in \widetilde{\mathcal{F}}_\delta} \left|\frac{1}{\prod_{k=1}^m \ell_k} \sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \left(\omega_{\overrightarrow{h}}^2\left(\hat{\mathbf{B}}_{i_1}^{(1)}, \ldots, \hat{\mathbf{B}}_{i_m}^{(m)}\right) - \mathbb{P}^m \omega_{\overrightarrow{h}}^2\right)\right| \to_p 0. \tag{46}$$

Verifying condition (45)

$$\max_{1 \leq j' \leq k} \mathbb{E}\left(\frac{\log \mathcal{N}\left(\delta, \mathcal{F}_M^2, e_{\ell, j'}\right)}{\ell_{j'}}\right)^{1/2}$$

$$\leq (\ell_1 \wedge \cdots \wedge \ell_k)^{-1/2} \mathbb{E}\left[\int_0^\delta \left(\log \mathcal{N}\left(\varepsilon, \mathcal{F}_M^2, e_{\ell, j'}\right)\right)^{m/2} \, d\varepsilon\right]$$

$$\leq \left(\frac{\delta}{\sqrt{2}M}\right)^{-1}(\ell_1 \wedge \cdots \wedge \ell_k)^{-1/2}\mathbb{E}\left[\int_0^{\delta/\sqrt{2}M}\left(\log \mathscr{N}\left(\varepsilon, \mathscr{F}_M^2, e_{\ell,j'}\right)\right)^{m/2}\mathrm{d}\varepsilon\right]$$

$$\leq (\delta/2M)^{-1}(\ell_1 \wedge \cdots \wedge \ell_k)^{-1/2}\int_0^1\left(\sup_Q \log \mathscr{N}\left(\varepsilon\|F\|_{L_2(Q)}, \mathscr{F}, L_2(Q)\right)\right)^{m/2}\mathrm{d}\varepsilon$$

$$\times \|F\|_{L_2(\mathbb{P}^m)} \to 0. \tag{47}$$

The shift from the second to the third line is true because

$$\mathscr{N}\left(\delta, \mathscr{F}_M^2, L_2(\mathbb{Q})\right) \leq \mathscr{N}\left(\delta/\sqrt{2}M, \mathscr{F}_M, L_2(\mathbb{Q})\right).$$

As the condition is verified, as well as $\ell_1 \wedge \cdots \wedge \ell_m \to \infty$, (46) follows directly using the previous proposition. Hence, there exists some sequence $\{a_\ell\}$, in a way that $a_\ell \to 0$ for any sequence $\{\delta_\ell\}$ with $\delta_\ell \to 0$ both under $\ell_1 \wedge \cdots \wedge \ell_m \to \infty$, such that:

$$\mathbb{E}\left\|\sum_{1\leq i_k\leq \ell_k, 1\leq k\leq m}\varepsilon_{i_1}^{(1)}\cdots\varepsilon_{i_m}^{(m)}\boldsymbol{\omega}_{\widetilde{h}}\left(\hat{\mathbf{B}}_{i_1}^{(1)}, \ldots, \hat{\mathbf{B}}_{i_m}^{(m)}\right)\right\|_{\widetilde{\mathscr{H}}} \leq a_\ell\left(\prod_{k=1}^m \ell_k\right)^{1/2}. \tag{48}$$

An application of Lemma 2 proves that

$$n^{-m/2}\mathbb{E}\left\|\sum_{1\leq i_1, \ldots, i_m\leq n}(\xi_{i_1}-1)\cdots(\xi_{i_m}-1)\boldsymbol{\omega}_{\widetilde{h}}(\hat{\mathbf{B}}_{i_1}, \ldots, \hat{\mathbf{B}}_{i_m})\right\|_{\mathscr{F}_{\delta_n}} \to 0, \quad n \to \infty.$$

This completes the proof for the asymptotic equicontinuity.

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.

Appendix A

This appendix contains supplementary information that is an essential part of providing a more comprehensive understanding of the paper. We also refer to [46] for more details.

Proof of Proposition 1. Let $\mathbf{B}_0 = \{X_1, \cdots, X_{\mathscr{T}_0}\}$ and $\mathbf{B}_{l_n}^{(n)} = \{X_{\mathscr{T}_{l_n-1}+1}, \ldots, X_n\}$ the possibly empty non-regenerative blocks of observations. Note that, for $l_n \leq 2$, the demonstration can be viewed directly in [59], under the assumptions *(C1), (C2)* and *(C3)*, we can see that

$$\mathbb{P}_\nu(l_n \leq 2) = O(n^{-2}).$$

Otherwise, for $l_n > 2$, we can write $W_n(h)$ as follows:

$$W_n(h) = (I) + (II),$$

where

$$
(I) = \frac{1}{\binom{n}{m}}\left[\sum_{1\le i_1<\ldots<i_{m-1}\le l_n-1}\omega_{\widetilde{h}}(\mathbf{B}_0,\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_{m-1}})\right.
$$

$$
\left.+\sum_{1\le i_1<\ldots i_{m-1}\le l_n-1}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_{m-1}},\mathbf{B}_{l_n})\right],
$$

$$
(II) = \frac{1}{\binom{n}{m}}\left[\sum_{j=2}^{m}\sum_{0\le k<i_1<\ldots i_{m-1-j}\le l_n}\omega_{\widetilde{h}}(\mathbf{B}_k,\ldots,\mathbf{B}_k,\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_{m-1-j}})\right.
$$

$$
\left.-\sum_{j=2}^{m}\sum_{1\le k<i_1<\ldots<i_{m-1-j}\le n}\widetilde{h}(X_k,\ldots,X_k,X_{i_1},\ldots,X_{i_m})\right]
$$

$$
= \frac{1}{\binom{n}{m}}\left[\sum_{\left(I_m^{l_n-1}\right)^c}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})-\sum_{(I_m^n)^c}\widetilde{h}(X_{i_1},\ldots,X_{i_m})\right],
$$

where

$$
(I_m^s)^c = \{(i_1,\ldots,i_m) : i_j\in\mathbb{N}, 1\le i_j\le n; \text{at least there are } j \text{ and } k \text{ such that } i_j = i_k\},
$$

the complement of index set, with cardinal equal to $\binom{s+m-1}{m}-\binom{s}{m}:=\overline{\binom{s}{m}}$. To prove the convergence of $W_n(h)$ to zero in probability, we must fulfill the convergence of (I) and (II) to zero in probability.

$$
\mathfrak{A} = \overline{\binom{n}{m}}^{-1}\sum_{\left(I_m^{l_n-1}\right)^c}\omega_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_m})
$$

$$
\underset{n\to\infty}{\to} \alpha^{-m}[\mathbb{E}(\omega_h(\mathbf{B}_1,\underbrace{\mathbf{B}_k,\ldots,\mathbf{B}_k}_{u\text{ times}},\mathbf{B}_{k+u},\ldots,\mathbf{B}_m))-\mathbb{E}_A((\tau)^u)(\mathbb{E}_A(\tau))^{m-u}\mu(h)],
$$

where $1\le k\le m$ and $1\le u\le k$. We apply the SLLN for Harris Markov chains to find the convergence of

$$
\mathfrak{B} = \overline{\binom{n}{m}}^{-1}\sum_{(I_m^n)^c}\widetilde{h}(X_{i_1},\ldots,X_{i_m}),
$$

to

$$
\int\ldots\int h(x_1,\underbrace{x_k,\ldots,x_k}_{u\text{ times}},x_{k+u},\ldots,x_m)d\mu(x_1)d\mu^u(x_1)d\mu(x_{k+u})\ldots d\mu(x_m)-\mu(h).
$$

Using the conditions, all terms in $\mathfrak{A}$ and $\mathfrak{B}$ are finite and we can prove the convergence of (II) to zero. Now, for (I), applying the SLLN and by Lemma 3.2 in [47] part i), we can see that

$$
\mathbb{P}_\nu\left(\lim_{n\to+\infty}\frac{l_n}{n}\to\alpha^{-1}=(\mathbb{E}_A(\tau))^{-1}\right)=1. \tag{A1}
$$

We have

$$
n^{-2m}\,\mathbb{E}_\nu\left[\left(\sum_{1\le i_1<\ldots<i_{m-1}\le l_n-1}\omega_{\widetilde{h}}(\mathbf{B}_0,\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_{m-1}})\right)^2\right]
$$

$$\leq \; 2\alpha^{-2m}\left\{\mathbb{E}_\nu\left[\left(\boldsymbol{\omega}_{|h|}(\mathbf{B}_0,\mathbf{B}_1,\ldots,\mathbf{B}_{m-1})^2\right] + \mathbb{E}_\nu(\boldsymbol{\tau}_0)(\mathbb{E}_A(\boldsymbol{\tau})^{m-1}\boldsymbol{\mu}(h)^2\right\} < \infty.$$

We obtain, in turn, that

$$n^{-2m}\mathbb{E}_\nu\left[\left(\sum_{1\leq i_1<\ldots i_{m-1}\leq l_n}\boldsymbol{\omega}_{\widetilde{h}}(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_{m-1}},\mathbf{B}_{l_n})\right)^2\right]$$

$$\leq \; 2\alpha^{-2m}\left\{\mathbb{E}_\nu\left[\left(\boldsymbol{\omega}_{|h|}(\mathbf{B}_1,\ldots,\mathbf{B}_{m-1},\mathbf{B}_{l_n})^2\right] + (\mathbb{E}_A(\boldsymbol{\tau}))^m\boldsymbol{\mu}(|h|)^2\right\} < \infty.$$

Hence, (I) also converges to zero a.s under $\mathbb{P}_\nu$ as $n \to \infty$. $\qquad\square$

Proof of Theorem 1. In what follows, let $L = l_n - 1$ denote the number of blocks observed. We find that

$$R_L(h) = S_L(h) + D_L(h),$$

where

$$S_L(h) \;=\; \frac{m}{L}\sum_{i=1}^{L}\widetilde{h}^{(1)}(\mathbf{B}_i),$$

$$D_L(h) \;=\; \sum_{j=2}^{m}\binom{m}{j}\binom{L}{j}^{-1}\sum_{1\leq i_1<\cdots<i_j\leq L}\widetilde{h}^{(j)}\left(\mathbf{B}_{i_1},\ldots,\mathbf{B}_{i_j}\right),$$

where $\widetilde{h}^{(c)}(\cdot)$ represents the conditional expectation of $\boldsymbol{\omega}_{\widetilde{h}}(\cdot)$ given the c of the coordinates, for all $\mathbf{B}_c \in \mathbb{T}$. The U-statistics $D_L(h)$ is obtained by truncating the Hoeffding decomposition after the first term $S_L(h)$. Then, we just need to show that:

1.
$$L^{1/2}S_L(h) \;\text{ converges weakly to a Gaussian process } \mathbb{G}_\mathbb{P} \text{ on } l^\infty(\mathscr{F}),$$

2.
$$\|L^{-m+1/2}D_L(h)\|_\mathscr{F} \to 0.$$

For $\mathbb{P}_L(h_1) := \frac{1}{L}\sum_{i=1}^{L}\widetilde{h}_1(\mathbf{B}_i)$, introduce

$$Z_L(h) = \sqrt{L}(\mathbb{P}_L(h_1) - \mathbb{P}(h_1)) = \frac{1}{\sqrt{L}}\sum_{i=1}^{L}\left(\mathbb{P}_L\left(\mathbb{P}^{m-1}(h)\right) - \mathbb{P}^m(h)\right).$$

Using (A1), we can replace the random variable $L = l_n - 1$ with the deterministic quantity $\check{L}$ and we write

$$Z_{\check{L}}(h) = \frac{1}{\sqrt{\check{L}}}\sum_{i=1}^{\check{L}}\left(\mathbb{P}_{\check{L}}\left(\mathbb{P}^{m-1}(h)\right) - \mathbb{P}^m(h)\right) + o_\mathbb{P},$$

where $\check{L} = 1 + \left\lfloor\frac{n}{\mathbb{E}_A(\boldsymbol{\tau})}\right\rfloor$. In order to establish the weak convergence for the empirical process $Z_{\check{L}}(h)$, it is sufficient and necessary to prove the finite dimensional convergence and the stochastic equicontinuity. For the finite multidimensional convergence, we have to prove that $\left(Z_{\check{L}}(h_{i_1}),\ldots,Z_{\check{L}}(h_{i_k})\right)$ converges weakly to $\left(G(h_{i_1}),\ldots,G(h_{i_k})\right)$, for every fixed finite collection of functions

$$\{h_{i_1},\ldots,h_{i_k}\} \subset \mathscr{F}.$$

In order to fix this, it is enough to show that for every fixed $a_1,\ldots,a_k \in \mathbb{R}$,

$$\sum_{j=1}^{k}a_j Z_{\check{L}}(h_{i_j}) \to N(0,\sigma^2), \text{ in distribution,}$$

where

$$\sigma^2 = \sum_{j=1}^{k} a_j^2 \text{Var}(Z_{\check{L}}(h_{i_j})) + \sum_{s \neq r} a_j a_i \text{Cov}(Z_{\check{L}}(h_{i_s}), Z_{\check{L}}(h_{i_r})).$$

By linearity, and in the same footsteps of the arguments of ([57], Chapter 17), we can prove that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{\check{L}} \widetilde{h}_1(\mathbf{B}_j) \to N\left(0, \gamma_{h_1}^2\right),$$

where, under Condition (C5),

$$\gamma_{h_1}^2 = \alpha \mathbb{E}_A\left(\widetilde{h}_1^2(\mathbf{B}_1)\right).$$

We readily infer that we have

$$\sqrt{L} S_L(h) \to N\left(0, m^2 \mathbb{E}_A\left(\widetilde{h}_1^2(\mathbf{B}_1)\right)\right).$$

Now, to verify the equicontinuity, we need to check that for every $\epsilon > 0$,

$$\lim_{\delta \to 0} \lim_{n \to \infty} \mathbb{P}\left(\sup_{d(f,g) \leq \delta} |Z_L(f) - Z_L(g)| > \epsilon\right) = 0,$$

where $d(\cdot, \cdot)$ is a pseudo distance for which the class $\mathscr{F}$ is totally bounded, and f, g belong to $\mathscr{F}$. According to [72], we have

$$
\begin{aligned}
|Z_L(f - g)| &= \left| \frac{1}{\sqrt{L}} \left[\sum_{k=1}^{L} (f-g)(\mathbf{B}_k) - \mathbb{P}^m(f-g) \right] \right| \\
&\leq \left| \frac{1}{\sqrt{L}} \sum_{a \leq k \leq b} ((f-g)(\mathbf{B}_k) - \mathbb{P}^m(f-g)) \right| \\
&\quad + \left| \frac{1}{\sqrt{L}} \sum_{1 \leq k \leq \lfloor n/\mathbb{E}(\tau) \rfloor} ((f-g)(\mathbf{B}_k) - \mathbb{P}^m(f-g)) \right|,
\end{aligned}
$$

where $a = \min(L, \lfloor n/\mathbb{E}(\tau) \rfloor)$ and $b = \max(L, \lfloor n/\mathbb{E}(\tau) \rfloor)$. For the left-hand part in the last inequality, we have

$$
\begin{aligned}
& \left| \sum_{a \leq k \leq b} ((f-g)(\mathbf{B}_k) - \mathbb{P}^m(f-g)) \right| \\
&\leq 2 \sup_{f \in \mathscr{F}} \max_{1 \leq s \leq n} \left\{ 2 \left| \sum_{1 \leq k \leq s} (f(\mathbf{B}_k) - \mathbb{P}^m(f)) \right| \right\}.
\end{aligned}
$$

Dividing the last inequality by $L^{1/2}$ and using the convergence result in ([72] Lemma 2.11) with Condition (C1), we obtain the desired result. The right-hand part in the inequality is treated using ([72] Lemma 4.2) providing that $\mathbb{E}_A(\tau)^{2+\alpha} < \infty$, for $\alpha > 0$ and the hypothesis of a finite uniform entropy integral. To complete the weak convergence of the regenerative U-statistic, we must treat the remaining terms of its Hoeffding decomposition. For $\zeta \in \mathscr{F}$, let us introduce

$$\zeta := \omega_{\widetilde{h}}(\mathbf{B}_1, \ldots, \mathbf{B}_m) - \mathbb{P}^m(h) - \sum_{i=1}^{m} \widetilde{h}^{(1)}(\mathbf{B}_i), \quad \mathbf{B}_i \in \mathbb{T}. \tag{A2}$$

Once can see that ζ is centered, that is

$$\int \zeta(\mathbf{B}_1, \ldots, \mathbf{B}_m) dP(\mathbf{B}_1) \ldots dP(\mathbf{B}_i) \ldots dP(\mathbf{B}_m) = 0. \tag{A3}$$

By the randomization theorem, according to [7] (for $r = 2$):

$$\mathbb{E}\left\|\sum_{1 \leq i_1 < \cdots < i_m \leq \check{L}} \zeta(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m})\right\| = \mathbb{E}\left\|\sum_{1 \leq i_1 < \cdots < i_m \leq \check{L}} \varepsilon_{i_1}^{(1)} \varepsilon_{i_2}^{(2)} \zeta(\mathbf{B}_{i_1}^{(1)}, \ldots, \mathbf{B}_{i_m}^{(1)})\right\|.$$

Hence, for $\mathfrak{C}$ a constant:

$$\mathbb{E}\left\|\check{L}^{-1/2} \sum_{1 \leq i_1 < \cdots < i_m \leq \check{L}} \zeta(\mathbf{B}_{i_1}, \ldots, \mathbf{B}_{i_m})\right\|_{\mathscr{F}} \leq \mathfrak{C} \mathbb{E} \int_0^\infty \check{L}^{-1/2} \log N_{n,2}(\varepsilon, \mathscr{F}) d\varepsilon.$$

It is sufficient now to use the theorem hypothesis of a uniform entropy integral to complete the proof of the theorem. $\square$

Proof of Theorem 2. We have

$$\mathbb{E}\left\|\sum_{1 \leq i_1 < \ldots < i_m \leq n} \xi_{i_1} \ldots \xi_{i_m} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\right\|_{\mathscr{F}}.$$

By decoupling of the U-process, due to [6],

$$\leq \mathbf{C}_m \, \mathbb{E}\left\|\sum_{1 \leq i_1 < \ldots < i_m \leq n} \xi_{i_1} \ldots \xi_{i_m} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\right\|_{\mathscr{F}}.$$

By symmetrization, due to [6], we have

$$\leq 2^m \mathbf{C}_m \, \mathbb{E}\left\|\sum_{1 \leq i_1 < \ldots < i_m \leq n} \xi_{i_1} \ldots \xi_{i_m} \varepsilon_{i_1}^{(1)} \ldots \varepsilon_{i_m}^{(m(n))} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\right\|_{\mathscr{F}}$$

for $(\mathrm{sgn}(\xi_1)\varepsilon_1^*, \ldots, \mathrm{sgn}(\xi_n)\varepsilon_n^*)$ a sequence independent and with the same distribution as $(\xi_1, \ldots, \xi_n)$. By the invariance of $(\mathbb{P}_\epsilon \otimes \mathbb{P})^{mn}$ and the fact that ξ is independent of $X^\cdot, \epsilon^\cdot$, we have that

$$= 2^m \mathbf{C}_m \, \mathbb{E}_{x,\epsilon}\left\|\sum_{1 \leq i_1 < \ldots < i_m \leq n} |\xi_{i_1}| \ldots |\xi_{i_m}| \mathrm{sgn}(\xi_{i_1})\varepsilon_{i_1}^{(1)} \ldots \mathrm{sgn}(\xi_{i_m})\varepsilon_{i_m}^{(m(n))} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\right\|_{\mathscr{F}}$$

$$= 2^m \mathbf{C}_m \, \mathbb{E}_{x,\epsilon}\left\|\sum_{1 \leq i_1 < \ldots < i_m \leq l_n - 1} |\xi_{i_1}| \ldots |\xi_{i_m}| \varepsilon_{i_1}^{(1)} \ldots \varepsilon_{i_m}^{(m(n))} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\right\|_{\mathscr{F}},$$

using the reversed order statistics of $\{|\xi_i|\}_{i=1}^n$, $|\xi_{(1)}| \geq \cdots \geq |\xi_{(n)}|$, and the permutations between the different sequences of random variables, and in the same footsteps as [46],

$$= 2^m \mathbf{C}_m \, \mathbb{E}\left\|\sum_{1 \leq i_1 < \ldots < i_m \leq n} |\xi_{(i_1)}| \ldots |\xi_{(i_m)}| \varepsilon_{i_1}^{(1)} \ldots \varepsilon_{i_m}^{(m(n))} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\right\|_{\mathscr{F}}$$

substituting $\xi_{(i)}$ by $\sum_{k=i}^n \xi_{(k)} - \xi_{(k+1)}$, with $|\xi_{(n+1)}| = 0$, we have

$$\leq 2^m \mathbf{C}_m \, \mathbb{E}\left\|\sum_{1 \leq i_1 < \ldots < i_m \leq n} \sum_{k_j \geq i_j, 1 \leq j \leq r} (|\xi_{(l_1)}| - |\xi_{(l_1+1)}|) \cdots (|\xi_{(l_m)}| - |\xi_{(l_m+1)}|)\right.$$
$$\left. \times \varepsilon_{i_1}^{(1)} \ldots \varepsilon_{i_m}^{(m(n))} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\right\|_{\mathscr{F}}$$

$$\leq 2^m \mathbf{C}_m \, \mathbb{E}\left\|\sum_{1 \leq i_1, \ldots, i_m \leq n} \sum_{\ell_k \geq i_k, 1 \leq k \leq m} (|\xi_{(\ell_1)}| - |\xi_{(\ell_1+1)}|) \cdots (|\xi_{(\ell_m)}| - |\xi_{(\ell_m+1)}|)\right.$$

$$
\times \epsilon_{i_1}^{(1)} \cdots \varepsilon_{i_m}^{(m(n))} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\Big\|_{\mathscr{F}}
$$

$$
\leq 2^m \mathbf{C}_m \ \mathbb{E}\left[\sum_{1 \leq \ell_1, \ldots, \ell_m \leq n}(|\xi_{(\ell_1)}| - |\xi_{(\ell_1+1)}|) \cdots (|\xi_{(\ell_m)}| - |\xi_{(\ell_m+1)}|)\right.
$$

$$
\left. \times \mathbb{E}\Big\|\sum_{1 \leq i_k \leq \ell_k, 1 \leq k \leq m} \epsilon_{i_1}^{(1)} \cdots \varepsilon_{i_m}^{(m(n))} f\left(Y_{i_1}^{(1)}, \ldots, Y_{i_m}^{(m)}\right)\Big\|_{\mathscr{F}}\right]
$$

$$
\leq 2^m \mathbf{C}_m \ \mathbb{E}\left[\sum_{1 \leq \ell_1, \ldots, \ell_m \leq n} \int_{|\xi_{(\ell_1+1)}|}^{|\xi_{(\ell_1)}|} \cdots \int_{|\xi_{(\ell_m+1)}|}^{|\xi_{(\ell_m)}|} \psi_n(\ell_1, \ldots, \ell_m) dt_m \cdots dt_1\right]
$$

$$
\leq 2^m \mathbf{C}_m \ \mathbb{E}\left[\sum_{1 \leq \ell_1, \ldots, \ell_m \leq n} \int_{|\xi_{(\ell_1+1)}|}^{|\xi_{(\ell_1)}|} \cdots \int_{|\xi_{(\ell_m+1)}|)}^{|\xi_{(\ell_m)}|}\right.
$$

$$
\left. \psi_n(|\{i : |\xi_i| > t_1\}|, \ldots, |\{i : |\xi_i| > t_m\}|) dt_m \cdots dt_1\right]
$$

$$
\leq 2^m \mathbf{C}_m \ \mathbb{E}\left[\int_{\mathbb{R}_{\geq 0}^m} \psi_n(|\{i : |\xi_i| > t_1\}|, \ldots, |\{i : |\xi_i| > t_m\}|) dt_1 \cdots dt_m\right]
$$

$$
\leq 2^m C_m \int_{\mathbb{R}_{\geq 0}^m} \mathbb{E}\psi_n\left(\sum_{i=1}^n \mathbb{1}_{|\xi_i| > t_1}, \ldots, \sum_{i=1}^n \mathbb{1}_{|\xi_i| > t_m}\right) dt_1 \cdots dt_m. \quad \text{(By Fubini's theorem.)}
$$

Now, suppose that $\psi_n(\ell_1, \ldots, \ell_m) = \bar{\psi}_n(\prod_{k=1}^m \ell_k)$. Then, we may further bound the above equation by

$$
\int_{\mathbb{R}_{\geq 0}^m} \mathbb{E}\bar{\psi}_n\left(\prod_{k=1}^m \sum_{i=1}^n \mathbb{1}_{|\xi_i| > t_k}\right) dt_1 \ldots dt_m
$$

$$
= \int_{\mathbb{R}_{\geq 0}^m} \mathbb{E}\bar{\psi}_n\left(\sum_{1 \leq i_1, \ldots, i_m \leq n} \prod_{k=1}^m \mathbb{1}_{|\xi_{i_k}| > t_k}\right) dt_1 \ldots dt_m
$$

$$
\leq \int_{\mathbb{R}_{\geq 0}^m} \bar{\psi}_n\left(\sum_{1 \leq i_1, \ldots, i_m \leq n} \mathbb{E}\prod_{k=1}^m \mathbb{1}_{|\xi_{i_k}| > t_k}\right) dt_1 \ldots dt_m \quad \text{(by Jensen's inequality)}
$$

$$
\leq \int_{\mathbb{R}_{\geq 0}^m} \bar{\psi}_n\left(\sum_{1 \leq i_1, \ldots, i_m \leq n} \prod_{k=1}^m \mathrm{P}\left(|\xi_{i_k}| > t_k\right)^{1/m}\right) dt_1 \ldots dt_m,
$$

where the last inequality follows from the generalized Hölder inequality and the assumption that $\bar{\psi}_n$ is nondecreasing. $\square$

Proof of Lemma 2. For

$$
\psi_n(\ell_1, \ldots, \ell_m) \equiv a(\ell_1, \ldots, \ell_m)\left(\prod_{k=1}^m \ell_k\right)^{1/2},
$$

Theorem 2 implies that:

$$
\mathbb{E}\Big\|\sum_{1 \leq i_1, \ldots, i_m \leq n} \xi_{i_1} \cdots \xi_{i_m} f\left(Y_{i_1}, \ldots, Y_{i_m}\right)\Big\|_{\mathscr{F}_{(n, \ldots, n), n}}
$$

$$
\leq K_m \int_{\mathbb{R}_{\geq 0}^m} \mathbb{E}\left[a\left(\sum_{i=1}^n \mathbf{1}_{|\xi_i| > t_1}, \ldots, \sum_{i=1}^n \mathbf{1}_{|\xi_i| > t_m}\right) \prod_{k=1}^m \left(\sum_{i=1}^n \mathbf{1}_{|\xi_i| > t_k}\right)^{1/2}\right] dt_1 \cdots dt_m
$$

$$
\leq K_m \int_{\mathbb{R}_{\geq 0}^m} A_{2,n}(t_1, \ldots, t_m)\left\{\mathbb{E}\prod_{k=1}^m \sum_{i=1}^n \mathbf{1}_{|\xi_i| > t_k}\right\}^{1/2} dt_1 \cdots dt_m
$$

$$\leq K_m \int_{\mathbb{R}^m_{\geq 0}} A_{2,n}(t_1, \ldots, t_m) \left(\sum_{1 \leq i_1, \ldots, i_m \leq n} \prod_{k=1}^m \mathbb{P}(|\xi_{i_k}| > t_k)^{1/m} \right)^{1/2} \mathrm{d}t_1 \cdots \mathrm{d}t_m$$

$$= n^{m/2} K_m \int_{\mathbb{R}^m_{\geq 0}} A_{2,n}(t_1, \ldots, t_m) \prod_{k=1}^m \mathbb{P}(|\xi_1| > t_k)^{1/2m} \, \mathrm{d}t_1 \cdots \mathrm{d}t_m.$$

Here

$$A_{2,n}(t_1, \ldots, t_m) \equiv \left\{ \mathbb{E} \left[a^2 \left(\sum_{i=1}^n \mathbf{1}_{|\xi_i| > t_1}, \ldots, \sum_{i=1}^n \mathbf{1}_{|\xi_i| > t_m} \right) \right] \right\}^{1/2} \to 0$$

as long as none of $\{\mathbb{P}(|\xi_1| > t_k) : 1 \leq k \leq m\}$ vanishes. The claim now follows from the dominated convergence theorem. $\square$

Corollary A1 ([6]). *Let $X(t), t \in T$, be a (weak) Gaussian or Rademacher chaos process of degree m and let*

$$d_X(s,t) := \left[\mathbb{E}(X(t) - X(s))^2 \right]^{1/2}, s, t \in T.$$

If

$$\int_0^D (\log N(T, d_X, \varepsilon))^{m/2} d\varepsilon < \infty,$$

then there is a version of X, which we keep denoting X, with almost all of its sample paths in $C_u(T, d_X)$ and such that

$$\left\| \sup_{t \in T} |X(t)| \right\|_{\psi_{2/m}} \leq \|X(t_0)\|_{\psi_2} + K \int_0^D (\log(N(T, d_X, \varepsilon)))^{m/2} d\varepsilon,$$

and

$$\left\| \sup_{\substack{d_X(s,t) \leq \delta \\ s,t \in T}} |X(t) - X(s)| \right\|_{\psi_{2/m}} \leq K \int_0^\delta (\log(N(T, d_X, \varepsilon)))^{m/2} d\varepsilon,$$

for all $0 < \delta \leq D$, where K is a universal constant and D is the diameter of T for the pseudodistance d_X. In fact, every separable version of X satisfies these properties.

Theorem A1 ([73]). *For any random elements Y_n with values in a metric space (S, d), where Y is measurable and has a separable range, the following are equivalent:*

1. *Y_n converge in law to Y;*
2. *$d_{\mathrm{BL}}(Y_n, Y) \to 0$ as $n \to \infty$;*

References

1. Halmos, P.R. The theory of unbiased estimation. *Ann. Math. Stat.* **1946**, *17*, 34–43. [CrossRef]
2. Hoeffding, W. A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **1948**, *19*, 293–325. [CrossRef]
3. van der Vaart, A.W. *Asymptotic Statistics*; Cambridge Series in Statistical and Probabilistic Mathematics; Cambridge University Press: Cambridge, MA, USA, 1998; Volume 3, p. xvi+443.
4. Mises, R.V. On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Stat.* **1947**, *18*, 309–348. [CrossRef]
5. Serfling, R.J. *Approximation Theorems of Mathematical Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2009; Volume 162.
6. 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; p. xvi+392.
7. Arcones, M.A.; Giné, E. Limit theorems for *U*-processes. *Ann. Probab.* **1993**, *21*, 1494–1542. [CrossRef]
8. Lee, A.J. *U-statistics*; Volume 110, *Statistics: Textbooks and Monographs*; Theory and practice; Marcel Dekker, Inc.: New York, NY, USA, 1990; p. xii+302.
9. Stute, W. Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **1993**, *21*, 146–156. [CrossRef]
10. Arcones, M.A.; Wang, Y. Some new tests for normality based on *U*-processes. *Statist. Probab. Lett.* **2006**, *76*, 69–82. [CrossRef]

11. Giné, E.; Mason, D.M. Laws of the iterated logarithm for the local U-statistic process. *J. Theoret. Probab.* **2007**, *20*, 457–485. [CrossRef]
12. 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]
13. Schick, A.; Wang, Y.; Wefelmeyer, W. Tests for normality based on density estimators of convolutions. *Statist. Probab. Lett.* **2011**, *81*, 337–343. [CrossRef]
14. Joly, E.; Lugosi, G. Robust estimation of *U*-statistics. *Stochastic Process. Appl.* **2016**, *126*, 3760–3773. [CrossRef]
15. Lee, S.; Linton, O.; Whang, Y.J. Testing for stochastic monotonicity. *Econometrica* **2009**, *77*, 585–602.
16. Ghosal, S.; Sen, A.; van der Vaart, A.W. Testing monotonicity of regression. *Ann. Statist.* **2000**, *28*, 1054–1082. [CrossRef]
17. Abrevaya, J.; Jiang, W. A nonparametric approach to measuring and testing curvature. *J. Bus. Econom. Statist.* **2005**, *23*, 1–19. [CrossRef]
18. Nolan, D.; Pollard, D. *U*-processes: Rates of convergence. *Ann. Statist.* **1987**, *15*, 780–799. [CrossRef]
19. Sherman, R.P. Maximal inequalities for degenerate *U*-processes with applications to optimization estimators. *Ann. Statist.* **1994**, *22*, 439–459. [CrossRef]
20. Yoshihara, K.i. Limiting behavior of *U*-statistics for stationary, absolutely regular processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **1976**, *35*, 237–252. [CrossRef]
21. 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]
22. Denker, M.; Keller, G. On *U*-statistics and v. Mises' statistics for weakly dependent processes. *Z. Wahrsch. Verw. Gebiete* **1983**, *64*, 505–522. [CrossRef]
23. Leucht, A. Degenerate *U*- and *V*-statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli* **2012**, *18*, 552–585. [CrossRef]
24. 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]
25. Bouzebda, S.; Nemouchi, B. Weak-convergence of empirical conditional processes and conditional *U*-processes involving functional mixing data. In *Statistical Inference for Stochastic Processes*; Springer: New York, NY, USA, 2022; pp. 1–56.
26. Efron, B. Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **1979**, *7*, 1–26. [CrossRef]
27. Hall, P. *The Bootstrap and Edgeworth Expansion*; Springer Series in Statistics; Springer: New York, NY, USA, 1992; p. xiv+352.
28. Bickel, P.J.; Freedman, D.A. Some asymptotic theory for the bootstrap. *Ann. Statist.* **1981**, *9*, 1196–1217. [CrossRef]
29. Arcones, M.A.; Giné, E. On the bootstrap of *U* and *V* statistics. *Ann. Statist.* **1992**, *20*, 655–674. [CrossRef]
30. Dehling, H.; Mikosch, T. Random quadratic forms and the bootstrap for *U*-statistics. *J. Multivariate Anal.* **1994**, *51*, 392–413. [CrossRef]
31. Leucht, A.; Neumann, M.H. Consistency of general bootstrap methods for degenerate *U*-type and *V*-type statistics. *J. Multivariate Anal.* **2009**, *100*, 1622–1633. [CrossRef]
32. Politis, D.N.; Romano, J.P. A circular block-resampling procedure for stationary data. In *Exploring the Limits of Bootstrap (East Lansing, MI, 1990)*; Wiley Ser. Probab. Math. Statist. Probab. Math. Statist.; Wiley: New York, NY, USA, 1992; pp. 263–270.
33. Carlstein, E. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **1986**, *14*, 1171–1179. [CrossRef]
34. Politis, D.N.; Romano, J.P. The stationary bootstrap. *J. Amer. Statist. Assoc.* **1994**, *89*, 1303–1313. [CrossRef]
35. Rubin, D.B. The Bayesian bootstrap. *Ann. Statist.* **1981**, *9*, 130–134. [CrossRef]
36. Lo, A.Y. A Bayesian method for weighted sampling. *Ann. Statist.* **1993**, *21*, 2138–2148. [CrossRef]
37. Mason, D.M.; Newton, M.A. A rank statistics approach to the consistency of a general bootstrap. *Ann. Statist.* **1992**, *20*, 1611–1624. [CrossRef]
38. Præstgaard, J.; Wellner, J.A. Exchangeably weighted bootstraps of the general empirical process. *Ann. Probab.* **1993**, *21*, 2053–2086. [CrossRef]
39. van der Vaart, A. New Donsker classes. *Ann. Probab.* **1996**, *24*, 2128–2140. [CrossRef]
40. Alvarez-Andrade, S.; Bouzebda, S. Strong approximations for weighted bootstrap of empirical and quantile processes with applications. *Stat. Methodol.* **2013**, *11*, 36–52. [CrossRef]
41. Bouzebda, S. On the strong approximation of bootstrapped empirical copula processes with applications. *Math. Methods Statist.* **2012**, *21*, 153–188. [CrossRef]
42. Bouzebda, S.; Elhattab, I.; Ferfache, A.A. General *M*-Estimator Processes and their *m* out of *n* Bootstrap with Functional Nuisance Parameters. In *Methodology and Computing in Applied Probability*; Springer: New York, NY, USA, 2022; pp. 1–45.
43. Huskova, M.; Janssen, P. Consistency of the generalized bootstrap for degenerate U-statistics. *Ann. Stat.* **1993**, *21*, 1811–1823. [CrossRef]
44. Janssen, P. Weighted bootstrapping of U-statistics. *J. Stat. Plan. Inference* **1994**, *38*, 31–41. [CrossRef]
45. Alvarez-Andrade, S.; Bouzebda, S. Cramér's type results for some bootstrapped *U*-statistics. *Statist. Papers* **2020**, *61*, 1685–1699. [CrossRef]
46. Han, Q. Multiplier *U*-processes: Sharp bounds and applications. *Bernoulli* **2022**, *28*, 87–124. [CrossRef]
47. Radulović, D. Renewal type bootstrap for Markov chains. *Test* **2004**, *13*, 147–192. [CrossRef]

48. Giné, E.; Zinn, J. Bootstrapping general empirical measures. *Ann. Probab.* **1990**, *18*, 851–869. [CrossRef]
49. Nummelin, E. *General Irreducible Markov Chains and Nonnegative Operators*; Cambridge Tracts in Mathematics; Cambridge University Press: Cambridge, UK, 1984; Volume 83, p. xi+156.
50. Athreya, K.B.; Ney, P. A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **1978**, *245*, 493–501. [CrossRef]
51. Nummelin, E. A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **1978**, *43*, 309–318. [CrossRef]
52. Chung, K.L. *Markov Chains with Stationary Transition Probabilities*, 2nd ed.; Die Grundlehren der mathematischen Wissenschaften, Band 104; Springer: New York, NY, USA, 1967; p. xi+301.
53. van der Vaart, A.W.; Wellner, J.A. *Weak Convergence and Empirical Processes*; Springer Series in Statistics; With applications to statistics; Springer: New York, NY, USA, 1996; p. xvi+508.
54. Pollard, D. *Convergence of Stochastic Processes*; Springer Series in Statistics; Springer: New York, NY, USA, 1984; p. xiv+215.
55. Douc, R.; Guillin, A.; Moulines, E. Bounds on regeneration times and limit theorems for subgeometric Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **2008**, *44*, 239–257. [CrossRef]
56. Meyn, S.; Tweedie, R.L. *Markov Chains and Stochastic Stability*, 2nd ed.; Cambridge University Press: Cambridge, MA, USA, 2009.
57. Meyn, S.P.; Tweedie, R.L. *Markov Chains and Stochastic Stability*; Communications and Control Engineering Series; Springer: London, UK, 1993; p. xvi+548.
58. Bertail, P.; Clémençon, S. Regeneration-based statistics for Harris recurrent Markov chains. In *Dependence in Probability and Statistics*; Springer: New York, NY, USA, 2006; Volume 187, pp. 3–54.
59. Bertail, P.; Clémençon, S. A renewal approach to Markovian U-statistics. *Math. Methods Statist.* **2011**, *20*, 79–105. [CrossRef]
60. Revuz, D. *Markov Chains*, 2nd ed.; North-Holland Publishing Co.: Amsterdam, The Netherlands, 1984; Volume 11.
61. Cheng, G. Moment consistency of the exchangeably weighted bootstrap for semiparametric M-estimation. *Scand. J. Stat.* **2015**, *42*, 665–684. [CrossRef]
62. Shao, J.; Wu, C. Heteroscedasticity-robustness of jackknife variance estimators in linear models. *Ann. Statist.* **1987**, *15*, 1563–1579. [CrossRef]
63. Weng, C.S. On a second-order asymptotic property of the Bayesian bootstrap mean. *Ann. Statist.* **1989**, *17*, 705–710. [CrossRef]
64. van Zwet, W.R. The Edgeworth expansion for linear combinations of uniform order statistics. In *Second Prague Symposium on Asymptotic Statistics (Hradec Králové, 1978)*; North-Holland: Amsterdam, NY, USA, 1979; pp. 93–101.
65. Pauly, M. Consistency of the subsample bootstrap empirical process. *Statistics* **2012**, *46*, 621–626. [CrossRef]
66. Shao, J.; Tu, D.S. *The Jackknife and Bootstrap*; Springer Series in Statistics; Springer: New York, NY, USA, 1995; p. xviii+516.
67. Bertail, P.; Clémençon, S. Regenerative block bootstrap for Markov chains. *Bernoulli* **2006**, *12*, 689–712. [CrossRef]
68. Bertail, P.; Clémençon, S. Approximate regenerative-block bootstrap for Markov chains. *Comput. Statist. Data Anal.* **2008**, *52*, 2739–2756. [CrossRef]
69. Fan, Y.; Ullah, A. On goodness-of-fit tests for weakly dependent processes using kernel method. *J. Nonparametr. Statist.* **1999**, *11*, 337–360. First NIU Symposium on Statistical Sciences (De Kalb, IL, 1996). [CrossRef]
70. Frees, E.W. Infinite order U-statistics. *Scand. J. Statist.* **1989**, *16*, 29–45.
71. Rempala, G.; Gupta, A. Weak limits of U-statistics of infinite order. *Random Oper. Stochastic Equations* **1999**, *7*, 39–52. [CrossRef]
72. Levental, S. Uniform limit theorems for Harris recurrent Markov chains. *Probab. Theory Related Fields* **1988**, *80*, 101–118. [CrossRef]
73. Dudley, R.M. Nonlinear functionals of empirical measures and the bootstrap. In *Probability in Banach Spaces, 7 (Oberwolfach, 1988)*; Progr. Probab.; Birkhäuser Boston: Boston, MA, USA, 1990; Volume 21, pp. 63–82.