Journal of Stochastic Analysis

Volume 1 | Number 3 Article 6

September 2020

Exchangeably Weighted Bootstraps of Martingale Difference Arrays under the Uniformly Integrable Entropy

Salim Bouzebda
Alliance Sorbonne Universités, Université de Technologie de Compiègne, L.M.A.C., Compiègne, France, salim.bouzebda@utc.fr

Nikolaos Limnios
LMAC, Universit´e de Technologie de Compi´egne, France, nikolaos.limnios@utc.fr

Journal of Stochastic Analysis
Vol. 1, No. 3 (2020) Article 6 (13 pages)
DOI: 10.31390/josa.1.3.06

EXCHANGEABLY WEIGHTED BOOTSTRAPS OF MARTINGALE DIFFERENCE ARRAYS UNDER THE UNIFORMLY INTEGRABLE ENTROPY

SALIM BOUZEBDA* AND NIKOLAOS LIMNIOS

ABSTRACT. In the present work, we are mainly concerned with the uniform central limit theorem for a bootstrapped martingale-difference array of a function-indexed stochastic process under the uniformly integrable entropy condition. More precisely, we establish the consistency of the exchangeable bootstraps.

1. Introduction and Motivation

The main idea of the present paper is to estimate the limiting distribution using the weighted bootstrap of sums of the form $\sum_{j\leq j(n)} V_{nj}(f)$, where the real-valued stochastic processes $\{V_{nj}(f) : f \in \mathcal{F}, 1 \leq j \leq j(n)\}$ for $n \geq 1$ are martingale-difference arrays on the probability space $(\Omega, \mathcal{E}, \mathbb{P})$ for arbitrary index set $\mathcal{F}$. It is worth nothing that the bootstrap was introduced and first investigated in Efron's seminal paper Efron, B. [14]. Since this seminal paper, bootstrap methods have been proposed, discussed, investigated and applied in an important number of papers in the scientific literature. Being one of the most important ideas in the applied statistics, the bootstrap also launched a wealth of innovative probability problems, which in turn formed the basis for the creation of new mathematical theories in probability and mathematical statistics. The main idea of the bootstrap is that if a sample is representative of the underlying population, then one can make inferences about the population characteristics by resampling from the current sample. Roughly speaking, it is known that the bootstrap works in the i.i.d. case if and only if the central limit theorem holds for the random variable under consideration. For further discussion, we refer the reader to the landmark paper Giné, E. and Zinn, J. [15]. Note that the limiting distributions of the processes that we are interested in, or their functionals, are rather complicated, which does not permit explicit computation in practice. More precisely, the limiting distributions of some statistics or functionals of the processes of interest depend in non trivial way of some unknown parameters, which can not be used in practical situation. The bootstrap methods are privileged alternatives to circumvent such difficulty. In the present work, we shall propose a general bootstrap methodology

Received 2020-2-29; Accepted 2020-7-19; Communicated by the editors.
2010 *Mathematics Subject Classification.* Primary 60F17; Secondary 60F05.
Key words and phrases. Uniform CLT, martingale difference array, uniformly integrable entropy, sequential empirical process, bootstrap.
* Corresponding author.

and study some of its asymptotic properties by means of the modern empirical processes theory. We extend our previous work Bouzebda, S. and Limnios, N. [9] by considering more general weight of the bootstraps.

The remainder of this paper is organized as follows. In Section 2, we provide some necessary background, state the functional central limit theorem that we are interested in, where the notation and definitions are consistent with the work of Bae, J., Jun, D., and Levental, S. [2]. The main theoretical results for the weighted bootstraps are given in Section 3. Some concluding remarks are given in Section 4. The proofs are given in the last Section 5.

2. Some Preliminaries and Notation

Bae, J., Jun, D., and Levental, S. [2] introduced the following general set up for studying the uniform central limit theorem for a specific function-indexed process based on a martingale-difference array under the assumption of uniformly integrable entropy. Let us consider an array of sub-σ-fields $\{\mathcal{E}_{nj} : 0 \leq j \leq j(n), n \in \mathbb{N}\}$ defined on some probability space $(\Omega, \mathcal{E}, \mathbb{P})$, that fulfills $\mathcal{E}_{n0} \subset \mathcal{E}_{n1} \subset \cdots \subset \mathcal{E}_{nj(n)}$ for $n \in \mathbb{N}$. For a set of real-valued functions $\mathcal{F}$ defined on a measurable space $(\mathbb{X}, \mathcal{X})$, let us introduce an array $\{V_{nj}(f) : j \leq j(n), n \in \mathbb{N}, f \in \mathcal{F}\}$ of martingale-difference of $\mathcal{L}_2$-process indexed by the set of functions $\mathcal{F}$ with respect to the σ-fields $\{\mathcal{E}_{nj} : 0 \leq j \leq n, n \in \mathbb{N}\}$, meaning that for any $f \in \mathcal{F}$, $\{V_{nj}(f) : j \leq j(n), n \in \mathbb{N}\}$ is an array of random variables with

$$\mathbb{E}(V_{nj}(f) \mid \mathcal{E}_{n,j-1}) = 0$$

and $V_{nj}(f)$ is $\mathcal{E}_{nj}$-measurable. Let us lighten our notation by writing $\mathbb{E}_{n,j-1}f$ to mean $\mathbb{E}(V_{nj}(f) \mid \mathcal{E}_{n,j-1})$, the usual conditional expectation of the random element $V_{nj}(f)$ given the σ-field $\mathcal{E}_{n,j-1}$. The conditional variance process is defined to be

$$v_{nj}(f) := \mathbb{E}_{n,j-1}(V_{nj}(f)^2)$$

for $f \in \mathcal{F}$. One can see that $v_{nj}(f)$ is also an $\mathcal{E}_{n,j-1}$-measurable random variable. For a class of measurable functions $\mathcal{F}$ defined on a measurable space $(\mathbb{X}, \mathcal{X})$, the covering number $N(\epsilon, \mathcal{F}, \|\cdot\|)$, denoted, if there is no ambiguity, by $N(\epsilon)$, is the minimum number of balls $\{g : \|g - h\| < \epsilon\}$ of radius ϵ needed to cover $\mathcal{F}$. We let F to be the envelope of the class of functions $\mathcal{F}$, meaning that F is a measurable function from $\mathbb{X}$ to $[0, \infty)$ satisfying

$$\sup_{f \in \mathcal{F}} |f(x)| \leq F(x), \text{ for all } x \in \mathbb{X}.$$

The set of all measures γ on $(\mathbb{X}, \mathcal{X})$ with

$$\gamma(F^2) := \int_{\mathbb{X}} F^2 d\gamma < \infty,$$

is denoted by $M(\mathbb{X}, F)$. For each (fixed or random) measures μ on $(\mathbb{X}, \mathcal{X})$, let us define

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

We use the notation $d := d_\gamma$, for $\gamma \in M(\mathbb{X}, F)$. One can say that the class of functions $\mathcal{F}$ has uniformly integrable entropy with respect to L_2-norm if the following condition is fulfilled

$$\int_0^\infty \sup_{\gamma \in M(\mathbb{X},F)} \left[\log\left(N\left(\epsilon[\gamma(F^2)]^{1/2}, \mathcal{F}, d_\gamma\right)\right)\right]^{1/2} d\epsilon < +\infty. \tag{2.1}$$

Notice that $(\mathcal{F}, d_\gamma)$ is totally bounded for any measure γ if the class $\mathcal{F}$ has uniformly integrable entropy. Many important classes of functions, such as VC graph classes, have uniformly integrable entropy. See Section 2.6 of van der Vaart and Wellner [30] and we may refer also to Kosorok [21]. In Theorem 1 in Chapter 8 of Pollard [26], among others, the one dimensional CLT for a martingale-difference array is given. In the sequel, the events are identified with their indicator functions and $\mathbb{E}^*$ denotes the upper expectation with respect to the outer probability $\mathbb{P}^*$. The aim here, is to estimate the limiting distribution of the process $\{S_n(f) : n \in \mathbb{N}, f \in \mathcal{F}\}$, which is given by

$$S_n(f) := \sum_{j \leq j(n)} V_{nj}(f), \text{ for } f \in \mathcal{F}.$$

Let us introduce

$$\sigma_n^2(f,g) := \sum_{j \leq j(n)} \mathbb{E}_{n,j-1}[V_{nj}(f) - V_{nj}(g)]^2, \text{ for all } f, g \in \mathcal{F}.$$

We now state two results of Bae, J., Jun, D., and Levental, S. [2] which is needed in the proof of our result.

Theorem 2.1. *([2]). Let $\{V_{nj}(f) : j \leq j(n), n \in \mathbb{N}, f \in \mathcal{F}\}$ be a martingale-difference array of $\mathcal{L}_2$-process indexed by a class of measurable functions $\mathcal{F}$ that admits the envelope function $F(\cdot)$ defined on a measurable space $(\mathbb{X}, \mathcal{X})$. Assume that the class of function $\mathcal{F}$ satisfy the condition (2.1). Denote by $\mu_n, n \in \mathbb{N}$, some random measures on $(\mathbb{X}, \mathcal{X})$ such that*

$$\mathbb{P}^* \left\{ \sup_{f,g \in \mathcal{F}} \frac{\sigma_n^2(f,g)}{(d_{\mu_n}(f,g))^2} \geq L \right\} \to 0, \text{ as } n \to \infty \text{ for a constant } L > 0. \tag{2.2}$$

Suppose

$$L_n(\sigma) := \frac{6}{\delta} \sum_{j \leq j(n)} \mathbb{E}[V_{nj}(F)^2 \{V_{nj}(F) > \delta\}] \to 0, \text{ for every } \delta > 0.$$

Then given $\epsilon > 0$ and $\gamma > 0$, there exists an $\eta > 0$ such that

$$\limsup_{n \to \infty} \mathbb{P}^* \left(\sup_{d_{\mu_n}(f,g) \leq \eta} |S_n(f) - S_n(g)| > 5\gamma \right) \leq 3\epsilon.$$

Notice that condition (2.2) is discussed in the paper by Bae, J., Jun, D., and Levental, S. [2].

Obtaining a uniform CLT essentially means proving that

$$\{\mathcal{L}(\mathbb{G}_n(f)) : f \in \mathcal{F}\} \to \{\mathcal{L}(Z(f)) : f \in \mathcal{F}\},$$

where the processes are indexed by $\mathcal{F}$ and are considered as random elements in the Banach space

$$B(\mathcal{F}) := \left\{ z : \mathcal{F} \to \mathbb{R} : \|z\|_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |z(f)| < \infty \right\},$$

the space of the bounded real-valued functions on $\mathcal{F}$, considered with respect to the sup norm. The limiting process $Z = \{Z(f) : f \in \mathcal{F}\}$ is a Gaussian process with sample paths are contained in

$$U_B(\mathcal{F}, \rho) := \{z \in B(\mathcal{F}) : z \text{ is uniformly continuous with respect to } \rho\},$$

where ρ is a metric on $\mathcal{F}$. By the fact that $U_B(\mathcal{F}, \rho)$ is a closed subspace of $(B(\mathcal{F}), \|\cdot\|_{\mathcal{F}})$ implies that it is also a Banach space. In particular $U_B(\mathcal{F}, \rho)$ is separable if and only if $(\mathcal{F}, \rho)$ is totally bounded. The space $\mathcal{F}$ will be equipped with the pseudometric d, so that $(\mathcal{F}, d)$ is totally bounded. Let us recall the following definition of weak convergence, introduced by Hoffmann-Jørgensen [16].

Definition 2.2. A sequence of $B(\mathcal{F})$-valued random functions $\{T_n : n \geq 1\}$ converges in law to a $B(\mathcal{F})$-valued Borel measurable random function T whose law concentrates on a separable subset of $B(\mathcal{F})$, denoted by $T_n \rightsquigarrow T$, if,

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

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

The second result we need from Bae, J., Jun, D., and Levental, S. [2] is the following theorem.

Theorem 2.3. *([2]). Let $\{V_{nj}(f) : j \leq j(n), n \in \mathbb{N}, f \in \mathcal{F}\}$ be an array of martingale-difference of $\mathcal{L}_2$-process indexed by a class measurable functions $\mathcal{F}$ with an envelope function F defined on a measurable space $(\mathbb{X}, \mathcal{X})$. Assume that the class of function $\mathcal{F}$ satisfy the condition (2.1). Assume that there exists a constant L such that*

$$\mathbb{P}^* \left\{ \sup_{f,g \in \mathcal{F}} \frac{\sigma_n^2(f,g)}{(d_{\mu_n}(f,g))^2} \geq L \right\} \to 0, \ \text{as } n \to \infty.$$

Suppose that, as $n \to \infty$, in probability

$$\sum_{j \leq j(n)} v_{nj}(f) \to \sigma^2(f), \ \text{for each } f \in \mathcal{F},$$

where $\sigma^2(f)$ are positive constants, and for every $\epsilon > 0$, in probability

$$\sum_{j \leq j(n)} \mathbb{E}_{n,j-1}((V_{nj}(F)^2)\mathbb{1}\{V_{nj}(F) > \epsilon\}) \to 0.$$

Suppose there exists a Gaussian process Z such that finite dimensional distributions of S_n converge to those of Z. Then

$$\{S_n(f), \ f \in \mathcal{F}\} \rightsquigarrow \{Z(f), f \in \mathcal{F}\} \ \text{as random elements of } B(\mathcal{F}).$$

The limiting process $Z = (Z(f) : f \in \mathcal{F})$ is mean zero Gaussian with covariance structure $\mathbb{E}Z(f)Z(g)$ and the sample paths of Z belong to $U_B(\mathcal{F}, d)$.

Bae, J., Jun, D., and Levental, S. [2] obtained, as a particular case of the last theorem, the uniform CLT for a sequence of martingale-difference in a previous paper Bae, J. and Choi, M. J. [1] with

$$V_{nj}(f) = n^{-1/2} D_j(f).$$

Let $\{D_j(f) : 1 \leq j \leq n, n \leq \mathbb{N}, f \in \mathcal{F}\}$ be a sequence of martingale-difference of $\mathcal{L}_2$-process indexed by a class $\mathcal{F}$ with respect to an increasing sequence of σ-fields $\{\mathcal{E}_j : 0 \leq j \leq n, n \in \mathbb{N}\}$. Assume that there exists a constant $L > 0$ such that

$$\mathbb{P}^* \left\{ \sup_{f,g \in \mathcal{F}} \sum_{j=1}^{n} \frac{\mathbb{E}_{n,j-1}[D_j(f) - D_j(g)]^2}{(nd(f,g))^2} \geq L \right\} \to 0, \text{ as } n \to \infty.$$

Suppose that, as $n \to \infty$, in probability

$$\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}_{j-1}(D_j(f))^2 \to \sigma^2(f), \text{ for each } f \in \mathcal{F},$$

where $\sigma^2(f)$ are positive constants, and for every $\epsilon > 0$, in probability

$$\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}_{j-1}((D_j(F)^2) \mathbb{1}\{D_j(F) > \epsilon\sqrt{n}\}) \to 0.$$

Suppose there exists a Gaussian process Z such that finite dimensional distributions of S_n converge to those of Z. Then

$$\left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^{n} D_j(f), \ f \in \mathcal{F} \right\} \rightsquigarrow \{Z(f), f \in \mathcal{F}\} \text{ as random elements of } B(\mathcal{F}).$$

3. Main Result

In this section, we shall establish the consistency of bootstrapping under general conditions in the framework of difference martingale arrays. Define, for each measurable function $f \in \mathcal{F}$,

$$\widehat{S}_n(f) := \{j(n)\}^{1/2} \sum_{j \leq j(n)} W_{ni} \left(V_{nj}(f, \omega) - \frac{1}{j(n)} \sum_{j \leq j(n)} V_{nj}(f, \omega) \right),$$

where W_{ni} are the bootstrap weights defined on the probability space $(\mathcal{W}, \Omega, \mathbb{P}_W)$. $V_{nj}(f, \omega)$ indicates that sequence $V_{nj}(f, \omega)$, for $j \leq j(n)$, is considered fixed. The bootstrap weights W_{ni}'s are assumed to belong to the class of exchangeable bootstrap weights introduced in Mason, D. M. and Newton, M. A. [23] and further investigated in Præstgaard, J. and Wellner, J. A. [27], Janssen, A. [18], Pauly, M. [24], Bouzebda, S. and Cherfi, M. [5] and Bouzebda, S. [4]. We shall assume the following conditions.

W.1 *The vector $W_n = (W_{n1}, \ldots, W_{nn})^\top$ is exchangeable for any $n = 1, 2, \ldots$, i.e., for any permutation $\pi = (\pi_1, \ldots, \pi_n)$ of $(1, \ldots, n)$, the joint distribution of*

$$\pi(W_n) = (W_{n\pi_1}, \ldots, W_{n\pi_n})^\top$$

is the same as that of W_n;

W.2

$$\max_{1\le i\le m_n} \left|W_{ni} - \overline{W}_n\right| = \max_{1\le i\le m_n} \left|W_{ni} - \frac{1}{m_n}\sum_{i=1}^{m_n} W_{ni}\right| \to 0, \text{ in probability,}$$

W.3

$$\sum_{i=1}^{m_n} \left(W_{ni} - \overline{W}_n\right)^2 \to 1, \text{ in probability,}$$

W.4

$$\sqrt{m_n}\left(W_{ni} - \overline{W}_n\right) \to Z, \text{ in distribution,}$$

where Z is a r.v. with $\mathbb{E}(Z) = 0$ and $Var(Z) = 1$.

In the usual Efron's nonparametric bootstrap, the bootstrap sample is drawn from the empirical distribution. One can show that $W_n \sim \text{Multinomial}(n; n^{-1}, \dots, n^{-1})$ satisfy the conditions W.1–W.4. The weights W_{ni} fulfil the conditions W.3-W.4 if some moment conditions are imposed, we may refer to Præstgaard, J. and Wellner, J. A. [27, Lemma 3.1]. The sampling schemes, *Bayesian bootstrap, Multiplier bootstrap, Double bootstrap, and Urn bootstrap*, fulfil the conditions W.1–W.4. These examples show that conditions W.1–W.4 are not restrictive. If the class $\mathcal{F}$ possesses enough measurability for randomization with i.i.d. multipliers to be possible, say a class of functions $\mathcal{F} \in M(\mathbb{P})$. It is woth noticing that $\mathcal{F} \in M(\mathbb{P})$, e.g., if $\mathcal{F}$ is countable, or if $\{S_n\}_n^\infty$ are stochastically separable in $\mathcal{F}$, or if $\mathcal{F}$ is image admissible Suslin; see Giné, E. and Zinn, J. [15, pages 853 and 854].
The main result of the present paper may now be stated precisely as follows.

Theorem 3.1. *We assume that* $\mathbf{W}$ *is a triangular array of bootstrap weights fulfilling assumptions W.1-W.4. Assume that the conditions of Theorem 2.3 hold. Then we have, almost surely,*

$$\{\widehat{S}_n(f),\ f \in \mathcal{F}\} \rightsquigarrow \{Z(f), f \in \mathcal{F}\} \text{ as random elements of } B(\mathcal{F}).$$

We regain the uniform CLT for a bootstrapped sequence of martingale-difference analogue to that in Bae, J. and Choi, M. J. [1] by applying Theorem 3.1 with $V_{nj}(f) = n^{-1/2}D_j(f)$ in a similar way as in Bae, J., Jun, D., and Levental, S. [2].

Corollary 3.2. *We assume that* $\mathbf{W}$ *is a triangular array of bootstrap weights fulfilling assumptions W.1-W.4. Assume that the conditions of Theorem 3.1 hold for* $\{D_j(f)/\sqrt{n} : f \in \mathcal{F}\}$*. Then we have, almost surely,*

$$\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n} W_{in}\left(D_j(f,w) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n} D_j(f,w)\right) : f \in \mathcal{F}\right\} \quad \rightsquigarrow \quad \{Z(f) : f \in \mathcal{F}\}.$$

3.1. Examples. Let us present some examples of the bootstrap weights satisfying the conditions W.1-W.4, we can refer to Præstgaard, J. and Wellner, J. A. [27] and Cheng, G. [11] for further details. More precisely, the following examples are provided in this compressed form in Cheng, G. [11], we have included some minor changes necessary for our setting. We may refer to Bouzebda, S. and Limnios, N. [6, 7, 8] and Bouzebda, S., Papamichail, Ch., and Limnios, N. [10].

Example 3.3 (i.i.d.-Weighted Bootstraps)**.** In this example, the bootstrap weights are defined as $W_{ni} = \omega_i/\overline{\omega}_n$, where $\omega_1, \omega_2, \ldots, \omega_n$ are i.i.d. positive r.v.s. with $\|\omega_1\|_{2,1} < \infty$, where

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

$$\overline{\omega}_n = \sum_{i=1}^n \omega_i.$$

Thus, we can choose $\omega_i \sim$ Exponential(1) or $\omega_i \sim$ Gamma(4, 1). The former corresponds to the Bayesian bootstrap. The multiplier bootstrap is often thought to be a smooth alternative to the nonparametric bootstrap; see Lo, A. Y. [22].

Example 3.4 (Efron's bootstrap)**.** As already mentioned, the weights for the Efron bootstrap satisfy the conditions W.1-W.5 with $c^2 = 1$ and are $W_n \sim$ Multinomial$(n; n^{-1}, \ldots, n^{-1})$.

Example 3.5 (The delete-h Jackknife)**.** In the delete-h jackknife, see Wu, C.F.J. [33], the bootstrap weights are generated by permuting the deterministic weights

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

Specifically, we have

$$W_{nj} = w_{nR_n(j)},$$

where $R_n(\cdot)$ is a random permutation uniformly distributed over $\{1, \ldots, n\}$. Thus, we need to choose $h/n \to \alpha \in (0, 1)$ such that $\varrho > 0$. Therefore, the usual jackknife with $h = 1$ is inconsistent for estimating the distribution.

Let us recall some examples from Janssen, A. [18].

Example 3.6. The $m(n)$ out of n-bootstrap weights

$$W_{ni} = m(n)^{1/2}\left(\frac{1}{m(n)}M_{ni} - \frac{1}{n}\right)$$

are given by a multinomial distributed random variable $(M_{n1}, \ldots, M_{n,n})$ with sample size

$$m(n) = \sum_{i=1}^n M_{ni}$$

and equal success probability. In this case, the conditions W.1-W.4 are valid, (details of the proof are given in Janssen, A. and Pauls, T. [19, (8.37)-(8.46)]).

Example 3.7. The $m(n)$-double bootstrap can be described by the weights

$$W_{ni} = \frac{m(n)^{1/2}}{\sqrt{2}}\left(\frac{1}{m(n)}M'_{ni} - \frac{1}{n}\right)$$

Here $(M'_{n1}, \ldots, M'_{nn})$ denotes a conditional multinomial distributed variable with sample size

$$m(n) = \sum_{i=1}^n M_{ni}$$

and success probability $M_{ni}/m(n)$ for the i-th cell given by the first example, (details of this example are discussed in Lemma 6.2 of Janssen, A. and Pauls, T. [19]).

Remark 3.8. Praestgaard and Wellner [27] pointed out that the weighted bootstraps are "smoother" in some sense than the multinomial bootstrap since they put some (random) weight at all elements in the sample, whereas the multinomial bootstrap puts positive weight at about

$$1 - (1 - n^{-1})^n \to 1 - e^{-1} = 0.6322$$

proportion of each element of the sample, on the average. Notice that when $\omega_i \sim \text{Gamma}(4, 1)$ so that the W_{ni}/n are equivalent to four-spacings from a sample of $4n - 1$ Uniform(0,1) random variables. In Weng [32] and Van Zwet [31], it was noticed that, in addition to being four times more expensive to implement, the choice of four-spacings crucially depends on the functional that we are interested in and is not universal.

3.2. Comments. Let us recall some comments from Bouzebda, S., Papamichail, Ch., and Limnios, N. [10]. Barbe, P. and Bertail, P. [3] discussed in details some properties of the weighted bootstrap for general von Mises functionals. The choice of the bootstrap weights depend on the applications at hand and the priorities of the statistician for specific situation : accuracy of the estimation of the entire distribution of the statistic; accuracy of a confidence interval related to coverage accuracy; accuracy in a large deviation sense; accuracy for a finite sample size. In the book of Shao, J. and Tu, D. S. [28] point out, for the bootstrap of the mean, that the random weighting method is less computationally intensive if n is not very large (this conclusion is in agreement with other references on the topic), on can refer to James, L. F. [17] and Shao, J. and Tu, D. S. [28], Tu, D. S. and Zheng, Z. G. [29], Chiang, C.-T., James, L. F., and Wang, M.-C. [13] and Chiang, C.-T., Wang, S.-H., and Hung, H. [12]. Finally, it is worth noticing that an appropriate choice of the bootstrap weights W_{ni}'s implies a smaller limit variance. For instance, typical example is the Subsample Bootstrap, Pauly [25, Remark 2.2-(3)].

4. Concluding Remarks

The functional central limit theorem of Bae, J., Jun, D., and Levental, S. [2] for martingale-difference array random processes has proven quite useful in establishing weak convergence results for several difficult statistical problems. A challenging task for doing inference in these settings is the fact that the limiting distributions are, in the most cases, very difficult to evaluate. To circumvent this, in the present paper, we have established the analogous results of Bae, J., Jun, D., and Levental, S. [2] for the bootstrap martingale-difference array random processes in an extended framework. Our results can be applied in the semi-Markov setting to construct confidence bands as Bouzebda, S., Papamichail, Ch., and Limnios, N. [10]. We mention also that the present paper largely extends the scope of applications of the last mentioned paper. The theoretical results established in this paper, are (or will be) key tools for many further developments in other settings.

5. Proof

This section is devoted to the proofs of our result. The aforementioned notation is also used in what follows. For a metric space $\{D, d\}$, let $BL_1(D)$ be the space of real-valued functions on D with Lipschitz norm bounded by 1, i.e., for any $f \in BL_1(D)$,

$$\sup_{x\in D} |f(x)| \leq 1$$

and

$$|f(x) - f(y)| \leq d(x, y)$$

for all $x, y \in D$.

Lemma 5.1. *Let $\{Y_{ni} : i = 1, \ldots, m_n, n \geq 1\}$ be a triangular array of mean zero real random vectors in $\mathbb{R}^d$ independent within rows; and let $\{W_{in} : i = 1, 2, \ldots, n, n = 1, 2, \ldots\}$ satisfy conditions W.1-W.4 and be independent of $\{Y_{ni} : i = 1, \ldots, m_n, n \geq 1\}$. Suppose also that*

(a) $\lim_{n\to\infty} \sum_{i=1}^{m_n} EY_{ni}Y_{ni}^\top = V_0 < \infty$ *where superscript* $\top$ *denotes transpose;*
(b) *for every* $\eta > 0$,

$$\limsup_{n\to\infty} \sum_{i=1}^{m_n} E\|Y_{ni}\|^2 \mathbb{1}\{\|Y_{ni}\| > \eta\} = 0,$$

where $\|\cdot\|$ *is the Euclidean norm on* $\mathbb{R}^d$.

Then

(i) $m_n^{1/2} \sum_{i=1}^{m_n} W_{ni}(Y_{ni} - \overline{Y}_{m_n})$ *converges weakly to* $Y_0 \overset{d}{=} N_d(0, V_0)$;
(ii)

$$\sup_{h\in BL_1(\mathbb{R}^d)} \left| E_W h\left(m_n^{1/2} \sum_{i=1}^{m_n} W_{ni}(Y_{ni} - \overline{Y}_{m_n}) \right) - E(h(Y_0)) \right| \to 0,$$

in probability, as $n \to \infty$, *where* $\mathbb{R}^d$ *is endowed with the uniform metric.*

Proof. Part (i) follows from Theorem 4.1 of Pauly, M. [24]. For part (ii) we use similar arguments those of Kosorok, M. R. [20]. Notice that (a) and (b) together imply that both, as $n \to \infty$, for every $\eta > 0$,

$$\sum_{i=1}^{m_n} \|Y_{ni}\|^2 \mathbb{1}\{Y_{ni} > \eta\} \to 0,$$

and

$$\sum_{i=1}^{m_n} Y_{ni}Y_{ni}^\top \to V_0.$$

Remark that

$$m_n^{1/2} \sum_{i=1}^{m_n} W_{ni}(Y_{ni} - \overline{Y}_{m_n}) = m_n^{1/2} \sum_{i=1}^{m_n} (W_{ni} - \overline{W}_{m_n})Y_{ni}.$$

This now implies that

$$\begin{aligned}
& E_W\left(m_n \sum_{i=1}^{m_n} W_{ni}^2 (Y_{ni}-\overline{Y}_{m_n})(Y_{ni}-\overline{Y}_{m_n})^\top\right) \\
&= E_W\left(m_n \sum_{i=1}^{m_n} (W_{ni}-\overline{W}_{m_n})^2 Y_{ni}Y_{ni}^\top\right) \\
&= \sum_{i=1}^{m_n} Y_{ni}Y_{ni}^\top E_W\left(m_n(W_{ni}-\overline{W}_{m_n})^2\right) \\
&= \sum_{i=1}^{m_n} Y_{ni}Y_{ni}^\top \to V_0
\end{aligned}$$

also converges in probability. Fix $\eta > 0$. Since

$$\begin{aligned}
& E_W\left(m_n \sum_{i=1}^{m_n} (W_{ni}-\overline{W}_{m_n})^2 Y_{ni}Y_{ni}^\top \mathbb{1}\left\{m_n^{1/2}|W_{ni}-\overline{W}_{m_n}|\|Y_{ni}\| > \eta\right\}\right) \\
\leq\ & E_W\left(m_n(W_{ni}-\overline{W}_{m_n})^2 \mathbb{1}\left\{m_n^{1/2}\left|W_{ni}-\overline{W}_{m_n}\right| > k\right\}\right)\sum_{i=1}^{m_n} Y_{ni}Y_{ni}^\top \\
& + k^2 \sum_{i=1}^{m_n} Y_{ni}Y_{ni}^\top \mathbb{1}\{\|Y_{ni}\| > \eta/k\},
\end{aligned}$$

for any positive $k < \infty$, we have that the left-hand side converges to zero in probability since we can choose k to make the first expectation on the right-hand side arbitrarily small. Since this is true for every $\eta > 0$, we can now replace η with a sequence $\{\eta_n\}$ going to zero. Thus for every subsequence n', there exists a further subsequence n'' such that

$$\begin{aligned}
\limsup_{n''\to\infty} E_W & \left(m_{n''} \sum_{i=1}^{m_{n''}} (W_{n''i}-\overline{W}_{m_{n''}})^2 Y_{n''i}Y_{n''i}^\top\right. \\
& \left.\mathbb{1}\left\{m_{n''}^{1/2}|W_{n''i}-\overline{W}_{m_{n''}}|\|Y_{n''i}\| > \eta\right\}\right) = 0,
\end{aligned}$$

almost surely. Thus, by the Lindeberg-Feller theorem combined with Theorem 1.12.2 of van der Vaart, A. W. and Wellner, J. A. [30], we have with probability 1 that

$$\lim n'' \to \infty \sup_{h\in BL_1(\mathbb{R}^d)} \left| E_W h\left(m_{n''}^{1/2} \sum_{i=1}^{m_{n''}} W_{n''i}(Y_{n''i}-\overline{Y}_{m_{n''}})\right) - E(h(Y_0))\right| = 0.$$

Since this is true for every subsequence n', part (ii) follows by Lemma 1.9.2 of van der Vaart, A. W. and Wellner, J. A. [30].

□

Proof. We will first apply Theorem 2.3 to

$$\left\{M_{nj} := V_{nj}(f,\omega)\left(W_{ni} - \frac{1}{j(n)}\sum_{j\leq j(n)} W_{ni}\right) : j \leq j(n), n\in\mathbb{N}, f\in\mathcal{F}\right\}.$$

For each $j \in \mathbb{N}$,

$$\mathbb{E}\left(M_{nj} \mid \mathcal{W} \times \mathcal{E}_{n,j-1}\right) = 0.$$

This, in turn, implies that $\{M_{nj} : j \leq j(n), n \in \mathbb{N}, f \in \mathcal{F}\}$-martingale difference for any $j \in \mathbb{N}$. Let us introduce, for $f \in \mathcal{F}$,

$$v^*_{nj}(f) := \mathbb{E}((M_{nj})^2 \mid \mathcal{W} \times \mathcal{E}_{n,j-1}) := \overline{\mathbb{E}}_{n,j-1}((M_{nj})^2).$$

First, we show that, as $n \to \infty$, in probability

$$\sum_{j \leq j(n)} \mathbb{E}_W v^*_{nj}(f) \to \sigma^2(f), \text{ for each } f \in \mathcal{F},$$

where $\sigma^2(f)$ are positive constants, and for every $\epsilon > 0$, in probability

$$\sum_{j \leq j(n)} \mathbb{E}_W \overline{\mathbb{E}}_{n,j-1}((M_{nj})^2 \mathbb{1}\{M_{nj} > \epsilon\}) \to 0.$$

Let

$$z_j = W_{ni} - \frac{1}{j(n)} \sum_{j \leq j(n)} W_{ni}.$$

We have

$$\begin{aligned}
&\sum_{j \leq j(n)} \mathbb{E}_W \mathbb{E}((z_j V_{nj}(f))^2 \mid \mathcal{W} \times \mathcal{E}_{n,j-1}) \\
&\quad = \sum_{j \leq j(n)} \mathbb{E}_W(z_j^2) \mathbb{E}((V_{nj}(f))^2 \mid \times \mathcal{E}_{n,j-1}) \\
&\quad = \sum_{j \leq j(n)} \mathbb{E}((V_{nj}(f))^2 \mid \times \mathcal{E}_{n,j-1}) \to \sigma^2(f).
\end{aligned}$$

Fix $\eta > 0$. Since

$$\begin{aligned}
&E_W \left(\sum_{i=1}^{m_n} \overline{\mathbb{E}}_{n,j-1}((z_j V_{nj}(F))^2 \mathbb{1}\{|z_j| V_{nj}(F) > \eta\}) \right) \\
&\leq \quad E_W \left(z_i^2 \mathbb{1}\left\{ |z_i| > k \right\} \right) \sum_{i=1}^{m_n} \mathbb{E}_{n,j-1}((V_{nj}(F))^2 \mathbb{1}\{V_{nj}(F) > \eta/k\}) \\
&\qquad + k^2 \sum_{i=1}^{m_n} \mathbb{E}_{n,j-1}((V_{nj}(F))^2 \mathbb{1}\{V_{nj}(F) > \eta/k\}),
\end{aligned}$$

for any positive $k < \infty$, we have that the left-hand side converges to zero in probability since we can choose k to make the first expectation on the right-hand side arbitrarily small. Let us define

$$\sigma_n^{2*}(f,g) := \sum_{j \leq j(n)} \overline{\mathbb{E}}_{n,j-1}[z_j V_{nj}(f) - z_j V_{nj}(g)]^2, \text{ for all } f, g \in \mathcal{F}.$$

For $L' = L_1 \times L > 0$, we want to show that

$$\mathbb{P}^* \left\{ \sup_{f,g \in \mathcal{F}} \frac{\sigma_n^{2*}(f,g)}{(d_{\mu_n}(f,g))^2} \geq L' \right\} \to 0, \text{ as } n \to \infty.$$

Notice that, under the Assumption W.2.,

$$\begin{aligned}
&\mathbb{P}^*\left\{\sup_{f,g\in\mathcal{F}}\frac{\sigma_n^{2*}(f,g)}{(d_{\mu_n}(f,g))^2}\geq L'\right\}\\
&\leq \mathbb{P}^*\left\{\left(\max_{j\leq j(n)} z_j\right)\sup_{f,g\in\mathcal{F}}\frac{\sigma_n^{2}(f,g)}{(d_{\mu_n}(f,g))^2}\geq L'\right\}\\
&\leq \mathbb{P}\left\{\max_{j\leq j(n)} z_j\geq L_1\right\}+\mathbb{P}^*\left\{\sup_{f,g\in\mathcal{F}}\frac{\sigma_n^{2}(f,g)}{(d_{\mu_n}(f,g))^2}\geq L\right\}\\
&= \mathbb{P}^*\left\{\sup_{f,g\in\mathcal{F}}\frac{\sigma_n^{2}(f,g)}{(d_{\mu_n}(f,g))^2}\geq L\right\}+o(1)\to 0,\ \text{as } n\to\infty.
\end{aligned}$$

Since this is true for every $\eta>0$, we can now replace η with a sequence $\{\eta_n\}$ going to zero. Thus, as in the proof of van der Vaart, A. W. and Wellner, J. A. [30, p. 357], for every subsequence n', there exists a further subsequence n'' such that

$$\limsup_{n''\to\infty} E_W\left(\sum_{i=1}^{m_{n''}}\overline{\mathbb{E}}_{n'',j-1}((z_jV_{n''j}(F))^2\mathbb{1}\{|z_j|V_{n''j}(F)>\eta\})\right)=0,$$

almost surely. From this point, the proof will follow the same line of Kosorok, M. R. [20], therefore is omitted. □

Acknowledgment. The authors are indebted to the referee for the very valuable comments and suggestions which led to a considerable improvement of this paper.

References

1. Bae, J. and Choi, M. J.: The uniform CLT for martingale difference of function-indexed process under uniformly integrable entropy. *Commun. Korean Math. Soc.*, **14**(1999), (3), 581–595.
2. Bae, J., Jun, D., and Levental, S.: The uniform CLT for martingale difference arrays under the uniformly integrable entropy. *Bull. Korean Math. Soc.*, **47**(2010) (1), 39–51.
3. Barbe, P. and Bertail, P.: *The weighted bootstrap*, volume 98 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1995.
4. Bouzebda, S.: On the strong approximation of bootstrapped empirical copula processes with applications. *Math. Methods Statist.*, **21**(2012) (3), 153–188.
5. Bouzebda, S. and Cherfi, M.: Test of symmetry based on copula function. *J. Statist. Plann. Inference*, **142**(2012) (5), 1262–1271.
6. Bouzebda, S. and Limnios, N.: On general bootstrap of empirical estimator of a semi-Markov Kernel with applications. *J. Multivariate Anal.*, **116** (2013a), 52–62.
7. Bouzebda, S. and Limnios, N.: Exchangeably weighted bootstraps of empirical estimators of a semi-Markov kernel. *C. R. Math. Acad. Sci. Paris*, **351**(2013b) (13-14), 569–573.
8. Bouzebda, S. and Limnios, N.: *Some Asymptotic Results for Exchangeably Weighted Bootstraps of the Empirical Estimator of a Semi-Markov Kernel with Applications*, In: *Statistical Models for Reliability and Survival Analysis*, Couallier V. et al. (Eds), Iste/J. Wiley, London, pp 211-226. 2013c.
9. Bouzebda, S. and Limnios, N.: Bootstraps of martingale difference arrays under the uniformly integrable entropy. In Statistical Topics and Stochastic Models for Dependant Data with Applications. (ISTE, Wiley). pp. 29–46. 2020.
10. Bouzebda, S., Papamichail, Ch., and Limnios, N.: On a multidimensional general bootstrap for empirical estimator of continuous-time semi-Markov kernels with applications. *J. Nonparametr. Stat.*, **30**(2018) (1), 49–86.

11. Cheng, G.: Moment Consistency of the Exchangeably Weighted Bootstrap for Semiparametric M-estimation. *Scand. J. Stat.* **42**(2015) (3), 665–684.
12. Chiang, C.-T., Wang, S.-H., and Hung, H.: Random weighting and edgeworth expansion for the nonparametric time-dependent AUC estimator. *Statist. Sinica*, **19**(2009) (3), 969–979.
13. Chiang, C.-T., James, L. F., and Wang, M.-C.: Random weighted bootstrap method for recurrent events with informative censoring. *Lifetime Data Anal.*, **11**(2005) (4), 489–509.
14. Efron, B.: Bootstrap methods: another look at the jackknife. *Ann. Statist.*, **7**(1979) (1), 1–26.
15. Giné, E. and Zinn, J.: Bootstrapping general empirical measures. *Ann. Probab.*, **18**(1990) (2), 851–869.
16. Hoffmann-Jørgensen, J.: *Stochastic processes on Polish spaces*, volume 39 of *Various Publications Series (Aarhus)*. Aarhus Universitet Matematisk Institut, Aarhus. 1991.
17. James, L. F.: A study of a class of weighted bootstraps for censored data. *Ann. Statist.*, **25**(1997) (4), 1595–1621.
18. Janssen, A.: Resampling Student's t-type statistics. *Ann. Inst. Statist. Math.*, **57**(2005) (3), 507–529.
19. Janssen, A. and Pauls, T.: How do bootstrap and permutation tests work? *Ann. Statist.*, **31**(2003) (3), 768–806.
20. Kosorok, M. R.: Bootstraps of sums of independent but not identically distributed stochastic processes. *J. Multivariate Anal.*, **84**(2003) (2), 299–318.
21. Kosorok, M. R.: *Introduction to empirical processes and semiparametric inference*. Springer Series in Statistics. Springer, New York. 2008.
22. Lo, A. Y.: A Bayesian method for weighted sampling. *Ann. Statist.*, **21**(1993) (4), 2138–2148.
23. Mason, D. M. and Newton, M. A.: A rank statistics approach to the consistency of a general bootstrap. *Ann. Statist.*, **20**(1992) (3), 1611–1624.
24. Pauly, M.: Weighted resampling of martingale difference arrays with applications. *Electron. J. Stat.*, **5**(2011), 41–52.
25. Pauly, M.: Consistency of the subsample bootstrap empirical process. *Statistics*, **46**(2012) (5), 621–626.
26. Pollard, D.: *Convergence of stochastic processes*. Springer Series in Statistics. Springer-Verlag, New York. 1984.
27. Præstgaard, J. and Wellner, J. A.: Exchangeably weighted bootstraps of the general empirical process. *Ann. Probab.*, **21**(1993) (4), 2053–2086.
28. Shao, J. and Tu, D. S.: *The jackknife and bootstrap*. Springer Series in Statistics. Springer-Verlag, New York. 1995.
29. Tu, D. S. and Zheng, Z. G.: Random weighting: another approach to approximate the unknown distributions of pivotal quantities. *J. Combin. Inform. System Sci.*, 16(2-3):249–270, 1991.
30. van der Vaart, A. W. and Wellner, J. A.: *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York. With applications to statistics. 1996.
31. van Zwet, W. R.: The Edgeworth expansion for linear combinations of uniform order statistics. In *Proceedings of the Second Prague Symposium on Asymptotic Statistics (Hradec Králové, 1978)*, pages 93–101. North-Holland, Amsterdam-New York. 1979.
32. Weng, C.-S.: On a second-order asymptotic property of the Bayesian bootstrap mean. *Ann. Statist.*, **17**(1989) (2), 705–710.
33. Wu, C.F.J.: On the asymptotic property of the jackknife histogram. *Technical report,* Dept. of Statistics, Univ. of Wisconsin, Madison. 1987.

Salim Bouzebda : LMAC, Université de Technologie de Compiégne, France
E-mail address: `salim.bouzebda@utc.fr`
URL: `https://bouzebda.github.io`

Nikolaos Limnios : LMAC, Université de Technologie de Compiégne, France
E-mail address: `nikolaos.limnios@utc.fr`