Electronic Journal of Statistics
ISSN: 1935-7524

Bootstrap uniform central limit theorems for Harris recurrent Markov chains *

Gabriela Ciołek

AGH University of Science and Technology,
al. Mickiewicza 30, 30-059 Krakow, Poland
Modal'X, Université Paris Ouest Nanterre la Défense,
200 Avenue de la République, 92000 Nanterre, France
e-mail: gabrielaciolek@gmail.com

Abstract: The main objective of this paper is to establish bootstrap uniform functional central limit theorem for Harris recurrent Markov chains over uniformly bounded classes of functions. We show that the result can be generalized also to the unbounded case. To avoid some complicated mixing conditions, we make use of the well-known regeneration properties of Markov chains. We show that in the atomic case the proof of the bootstrap uniform central limit theorem for Markov chains for functions dominated by a function in L^2 space proposed by Radulović (2004) can be significantly simplified. Finally, we prove bootstrap uniform central limit theorems for Fréchet differentiable functionals in a Markovian setting.

Primary 62G09; secondary 62G20, 60J05.
Keywords and phrases: Bootstrap, Markov chains, regenerative processes, Nummelin splitting technique, empirical processes indexed by classes of functions, entropy, robustness, Fréchet differentiability.

1. Introduction

The naive bootstrap for indentically distributed and independent random variables introduced by Efron (1979) has gradually evolved and new types of bootstrap schemes in both i.i.d. and dependent setting were established. A detailed review of various bootstrap methods such as moving block bootstrap (MBB), nonoverlapping block bootstrap (NBB) or cilcular block bootstrap (SBB) for dependent data can be found in Lahiri (2003). The main idea of block bootstrap procedures is to resample blocks of observations in order to capture the dependence structure of the original sample. However, as indicated by many authors, these procedures struggle with many problems. For instance, popular MBB method requires the stationarity for observations that usually results in failure of this method in non-stationary setting (see Lahiri (2003) for more details). Furthermore, the asymptotic behaviour of MBB method is highly dependent on the estimation of the bias and of the asymptotic variance of the statistic of interest that is a significant drawback when considering practical applications.

*Author is a beneficiary of the French Government scholarship Bourse Eiffel, managed by Campus France.

Finally, it is noteworthy, that the rate of convergence of the MBB distribution is slower than that's of bootstrap distribution in the i.i.d. setting. Moreover, all mentioned block bootstrap procedures struggle with the problem of the choice of the length of the blocks of data in order to reflect the dependence structure of the original sample.

It is rather surprising that the bootstrap theory for Markov chains has been paid relatively limited attention given the extensive investigation and development of various bootstrap methods for both i.i.d. and dependent data. One of the first bootstrap results for Markov chains was obtained by Datta and McCormick (1993). The proposed method relies on the renewal properties of Markov chains when a (recurrent) state is visited infinitely often. The idea behind the procedure is to resample a deterministic number of data blocks which are corresponding to regeneration cycles. However, the method proposed by Datta and McCormick is not second-order correct. Bertail and Clémençon (2007) have proposed the modification of this procedure which gives the second-order correctness in the stationary case, but fails in the nonstationary setting. Bertail and Clémençon (2006) have proposed two effective methods for bootstrapping Markov chains: the Regenerative block bootstrap (RBB) method for atomic chains and Approximate block bootstrap method (ARBB) for general Harris recurrent Markov chains. The main idea behind these procedures is to mimick the renewal (pseudo-renewal in general Harris case) structure of the chain by drawing regeneration data blocks, until the length of the reconstructed bootstrap sample is larger than the length of the original data. Blocks before the first and after the last regeneration times are discarded in order to avoid large bias. In the atomic setting, the RBB method has the uniform rate of convergence of order $O_{\mathbb{P}}(n^{-1})$ which is the optimal rate of convergence in the i.i.d. case. Bertail and Clémençon (2006) have proved the second-order correctness of the ARBB procedure in the unstudentized stationary case, the rate of convergence is close to that in the i.i.d. case. It is noteworthy that for both methods, the division of the data into blocks is completely data-driven what is a significant advantage in comparison to block bootstrap methods. It is worthy of mention, that in parallel to the paper of Bertail and Clémençon (2006), the Markov chains bootstrap CLT for the mean under no additional assumptions was proposed by Radulović (2004).

Bootstrap results for Markov chains established by Radulović (2004) and Bertail and Clémençon (2006) allow naturally to extend the bootstrap theory to empirical processes indexed by classes of functions in a Markovian setting. Radulović (2004) has proved the bootstrap uniform functional central limit theorem over uniformly bounded classes of functions $\mathcal{F}$. In mentioned paper, Radulović considers countable regenerative Markov chains and indicates that with additional uniform entropy condition the bootstrap result can be extended to the uncountable case. Gorst-Rasmussen and Bøgsted (2009) have proved the bootstrap uniform central limit theorem over classes of functions whose envelope is in L^2. They have considered regenerative case which was motivated by their study of queuing systems with abandonment.

This paper generalizes the Radulović's (2004) bootstrap result for empirical processes for Markov chains. We establish the bootstrap uniform functional central limit theorem over a permissible uniformly bounded classes of functions in general Harris case. We also show that by arguments of Tsai (1998), the condition of the uniform boundedness of $\mathcal{F}$ can be weakened and it is sufficient to require only that $\mathcal{F}$ has an envelope F in L^2. The proof of the bootstrap uniform CLT for Harris recurrent Markov chains is closely related to the uniform CLT for countable atomic Markov chains proposed by Radulović. Similarly as in his paper, the main struggle is the random number of pseudo-regeneration blocks. However, using regeneration properties of Markov chains, it is possible to replace the random number of blocks with its deterministic equivalent what simplifies the analysis of asymptotic properties of the studied empirical process. The arguments from the proof of main theorem of this paper can be also applied directly to the proof of bootstrap uniform CLT for atomic Markov chains proposed by Radulović (2004). Thus, we can significantly simplify the proof of the Radulović's result and apply standard probability inequalities for i.i.d. blocks of data to show the asymptotic stochastic equicontinuity of the bootstrap version of original empirical process indexed by uniformly bounded class of function.

Regenerative properties of Markov chains can be applied in order to extend some concepts in robust statistics from i.i.d. to a Markovian setting. Martin and Yohai (1986) have shown that, generally, proving that statistics are robust in dependent case is a challanging task. Bertail and Clémençon (2006) have defined an influence function and Fréchet differentiability on the torus what allowed to extend the notion of robustness from single observations to the blocks of data instead. As shown in Bertail and Clémençon (2015), this approach leads directly to central limit theorems (and their bootstrap versions) for Fréchet differentiable functionals in a Markovian setting. In our framework, we use the bootstrap asymptotic results for empirical processes indexed by classes of functions to derive bootstrap uniform central limit theorems for Fréchet differentiable functionals in a Markovian case. Interestingly, there is no need to consider blocks of data as in Bertail and Clémençon (2015). We show that the theorems work when classes of functions are permissible and uniformly bounded, however, it is easy to weaken the last assumption and impose that $\mathcal{F}$ has an envelope in L^2.

The paper is organized as follows. In section 2, we introduce the notation and preliminary assumptions for Markov chains. In section 3, we recall briefly some bootstrap methods for Harris recurrent Markov chains and formulate further necessary assumptions for the considered Markov chains. In section 4, we establish the bootstrap uniform central limit theorem for Markov chains. We give a proof for uniformly bounded classes of functions and show how the theory can be easily extended to the unbounded case. We indicate that using regeneration properties of Markov chains, the proof of uniform bootstrap central limit theorem for countable chains proposed by Radulović can be simplified. In section 5, the bootstrap uniform central limit theorems for Fréchet differentiable functionals in a Markovian setting are established. We prove that the central limit theorem holds when classes of functions are uniformly bounded. Next, we generalize the theory to the unbounded case demanding that $\mathcal{F}$ has an envelope in L^2. In the last section, we enclose small appendix with a proof of the interesting property used in the proofs of main asymptotic theorems in the previous section.

2. Preliminaries

We begin by introducing some notation and recall the key concepts of the Markov chains theory (see Meyn & Tweedie (1996) for a detailed review and references). For the reader's convenience we keep our notation in agreement with notation set in Bertail and Clémençon (2006). All along this section $\mathbb{I}_A$ is the indicator function of the event A.

Let $X = (X_n)_{n\in\mathbb{N}}$ be a homogeneous Markov chain on a countably generated state space $(E, \mathcal{E})$ with transition probability Π and initial probability ν. Note that for any $B \in \mathcal{E}$ and $n \in \mathbb{N}$, we have

$$X_0 \sim \nu \ \text{ and } \ \mathbb{P}(X_{n+1} \in B | X_0, \cdots, X_n) = \Pi(X_n, B) \ \text{ a.s.}$$

In our framework, $\mathbb{P}_x$ (resp. $\mathbb{P}_\nu$) denotes the probability measure such that $X_0 = x$ and $X_0 \in E$ (resp. $X_0 \sim \nu$), and $\mathbb{E}_x(\cdot)$ is the $\mathbb{P}_x$-expectation (resp. $\mathbb{E}_\nu(\cdot)$ is the $\mathbb{P}_\nu$-expectation). In the following, we assume that X is ψ -irreducible and aperiodic, unless it is specified otherwise.

We are particularly interested in the atomic structure of Markov chains. It is shown by Nummelin (1978) that any chain that possesses some recurrent properties can be extended to a chain which has an atom.

Definition 2.1. *Assume that X is aperiodic and ψ-irreducible. We say that a set $A \in \mathcal{E}$ is an accessible atom if for all $x,\ y \in A$ we have $\Pi(x, \cdot) = \Pi(y, \cdot)$ and $\psi(A) > 0$. In that case we call X atomic.*

In our framework, we are interested in the asymptotic behaviour of positive recurrent Harris Markov chains. We say that X is *Harris recurrent* if starting from any point $x \in E$ and any set such that $\psi(A) > 0$, we have $\mathbb{P}_x(\tau_A < +\infty) = 1$. Observe that the property of Harris recurrence ensures that X visits set A infinitely often a.s.. It follows directly from the strong Markov property, that given any initial law ν, the sample paths can be divided into i.i.d. blocks corresponding to the consecutive visitis of the chain to atom A. The segments of data are of the form:

$$\mathcal{B}_j = (X_{1+\tau_A(j)}, \cdots, X_{\tau_A(j+1)}),\ j \geq 1$$

and take values in the torus $\cup_{k=1}^\infty E^k$.

We define the sequence of regeneration times $(\tau_A(j))_{j\geq 1}$. The sequence consists of the successive points of time when the chain forgets its past. Let

$$\tau_A = \tau_A(1) = \inf\{n \geq 1 : X_n \in A\}$$

be the first time when the chain hits the regeneration set A and

$$\tau_A(j) = \inf\{n > \tau_A(j-1), X_n \in A\} \text{ for } j \geq 2.$$

In our framework, we consider steady-state behaviour of Markov chains. One of the crucial stability results of interest is the Kac's theorem which enables to write functionals of the stationary distribution μ as the functionals of distribution of a regenerative block. Indeed, for positive recurrent Markov chain if $\mathbb{E}_A(\tau_A) < \infty$, then the unique invariant probability distribution μ is the Pitman's occupation measure given by

$$\mu(B) = \frac{1}{\mathbb{E}_A(\tau_A)} \left(\sum_{i=1}^{\tau_A} \mathbb{I}\{X_i \in B\} \right) \ \forall B \in \mathcal{E}.$$

We introduce few more pieces of notation: throughout the paper we write $l_n = \sum_{i=1}^{n} \mathbb{I}\{X_i \in A\}$ for the total number of consecutive visits of the chain to the atom A, thus we observe l_n+1 data blocks. We make the convention that $B_{l_n}^{(n)} = \emptyset$ when $\tau_A(l_n) = n$. Furthermore, we denote by $l(B_j) = \tau_A(j+1) - \tau_A(j),\ j \geq 1$, the length of regeneration blocks. Note that the by the Kac's theorem we have that $\mathbb{E}(l(B_j)) = \mathbb{E}_A(\tau_A) = \frac{1}{\mu(A)}$. Consider $\mu-$ integrable function $f:\ E \to \mathbb{R}$. By $u_n(f) = \frac{1}{\tau_A(l_n) - \tau_A(1)} \sum_{i=1}^{n} f(X_i)$ we denote the estimator of the unknown asymptotic mean $\mathbb{E}_\mu(f(X_1))$.

Remark 2.1. *In order to avoid large bias of the estimators based on the regenerative blocks we discard the data before the first and after the last pseudo-regeneration times (for more details refer to Bertail and Clémençon (2006), page 693).*

2.1. General Harris Markov chains and the splitting technique

In this subsection, we recall the so-called *splitting technique* introduced in Nummelin (1978). The technique allows to extend the probabilistic structure of any Harris chain in order to artificially construct a regeneration set. In the following, unless specified otherwise, X is a general, aperiodic, ψ-irreducible chain with transition kernel Π.

Definition 2.2. *We say that a set $S \in \mathcal{E}$ is small if there exists a parameter $\delta > 0$, a positive probability measure Φ supported by S and an integer $m \in N^*$ such that*

$$\forall x \in S,\ A \in \mathcal{E} \ \ \Pi^m(x, A) \geq \delta\ \Phi(A), \tag{2.1}$$

where Π^m denotes the m-th iterate of the transition probability Π.

Remark 2.2. *It is noteworthy that in general case it is not obvious that small sets having positive irreducible measure exist. Jain and Jamison (1967) showed that they do exist for any irreducible kernel Π under the assumption that the state space is countably generated.*

We expand the sample space in order to define a sequence $(Y_n)_{n\in\mathbb{N}}$ of independent r.v.'s with parameter δ. We define a joint distribution $\mathbb{P}_{\nu,\mathcal{M}}$ of $X^{\mathcal{M}} = (X_n, Y_n)_{n\in\mathbb{N}}$. The construction relies on the mixture representation of Π on

S, namely $\Pi(x,A) = \delta\Phi(A) + (1-\delta)\frac{\Pi(x,A)-\delta\Phi(A)}{1-\delta}$. It can be retrieved by the following randomization of the transition probability Π each time the chain X visits the set S. If $X_n \in S$ and

- if $Y_n = 1$ (which happens with probability $\delta \in]0,1[$), then X_{n+1} is distributed according to the probability measure Φ,
- if $Y_n = 0$ (that happens with probability $1-\delta$), then X_{n+1} is distributed according to the probability measure $(1-\delta)^{-1}(\Pi(X_n,\cdot) - \delta\Phi(\cdot))$.

This bivariate Markov chain $X^{\mathcal{M}}$ is called the *split chain*. It takes its values in $E \times \{0,1\}$ and possesses an atom, namely $S \times \{1\}$. The split chain $X^{\mathcal{M}}$ inherits all the stability and communication properties of the chain X. The regenerative blocks of the split chain are i.i.d. (in case m = 1 in (2.1)). If the chain X satisfies $\mathcal{M}(m,S,\delta,\Phi)$ for $m > 1$, then the blocks of data are 1-dependent, however, it is easy to adapt the theory from the case when $m = 1$ (see for instance Levental (1988)).

2.2. Regenerative blocks for dominated families

Throughout the rest of the paper, the minorization condition $\mathcal{M}$ is fulfilled with $m = 1$, unless specified otherwise. We assume that the family of the conditional distributions $\{\Pi(x,dy)\}_{x\in E}$ and the initial distribution ν are dominated by a σ-finite measure λ of reference, so that $\nu(dy) = f(y)\lambda(dy)$ and $\Pi(x,dy) = p(x,y)\lambda(dy)$, for all $x \in E$. The minorization condition requests that Φ is absolutely continuous with respect to λ and that $p(x,y) \geq \delta\phi(y)$, $\lambda(dy)$ a.s. for any $x \in S$, with $\Phi(dy) = \phi(y)dy$. Consider the binary random sequence Y constructed via the Nummelin's technique from the parameters inherited from condition $\mathcal{M}$. We want to approximate the Nummelin's construction. Note that the distribution of $Y^{(n)} = (Y_1, ..., Y_n)$ conditionally to $X^{(n+1)} = (x_1, ..., x_{n+1})$ is the tensor product of Bernoulli distributions given by: for all $\beta^{(n)} = (\beta_1, ..., \beta_n) \in \{0,1\}^n$, $x^{(n+1)} = (x_1, ..., x_{n+1}) \in E^{n+1}$,

$$\mathbb{P}_\nu\left(Y^{(n)} = \beta^{(n)} \mid X^{(n+1)} = x^{(n+1)}\right) = \prod_{i=1}^{n} \mathbb{P}_\nu(Y_i = \beta_i \mid X_i = x_i,\ X_{i+1} = x_{i+1}),$$

with, for $1 \leqslant i \leqslant n$,

- if $x_i \notin S$, $\mathbb{P}_\nu(Y_i = 1 \mid X_i = x_i,\ X_{i+1} = x_{i+1}) = \delta$,
- if $x_i \in S$, $\mathbb{P}_\nu(Y_i = 1 \mid X_i = x_i,\ X_{i+1} = x_{i+1}) = \delta\phi(x_{i+1})/p(x_i, x_{i+1})$.

Observe that conditioned on $X^{(n+1)}$, from $i = 1$ to n, Y_i is distributed according to the Bernoulli distribution with parameter δ, unless X has hit the small set S at time i: then, Y_i is drawn from the Bernoulli distribution with parameter $\delta\phi(X_{i+1})/p(X_i, X_{i+1})$. We denote by $\mathcal{L}^{(n)}(p, S, \delta, \phi, x^{(n+1)})$ this probability distribution. If we were able to generate $Y_1, ..., Y_n$, so that $X^{\mathcal{M}(n)} = ((X_1, Y_1), ..., (X_n, Y_n))$ be a realization of the split chain $X^{\mathcal{M}}$, then we would be able to do the block decomposition of the sample path $X^{\mathcal{M}(n)}$ leading to asymptotically i.i.d. blocks. Note, that in the above procedure the knowledge about the transition density $p(x, y)$ is required in order to generate random variables $(Y_1, \cdots, Y_n)$. To deal with this problem in practice, Bertail and Clémençon (2006) proposed the approximating construction of the above procedure. We proceed as follows. We construct an estimator $p_n(x, y)$ of $p(x, y)$ based on $X^{(n+1)}$ (and $p_n(x, y)$ satisfies $p_n(x, y) \geq \delta\phi(y)$, $\lambda(dy)-$a.s. and $p_n(x, y) > 0, 1 \leq i \leq n$). Next, we generate random vector $\hat{\mathbb{Y}}_n = (\hat{Y}_1, \cdots, \hat{Y}_n)$ conditionally to $X^{(n+1)}$ from distribution $\mathcal{L}^{(n)}(p_n, S, \delta, \gamma, X^{(n+1)})$ which is an approximation of the conditional distribution $\mathcal{L}^{(n)}(p, S, \delta, \gamma, X^{(n+1)})$ of $(Y_1, \cdots, Y_n)$ for given $X^{(n+1)}$.

In this setting, we define the successive hitting times of $A_{\mathcal{M}} = S \times \{1\}$ as $\hat{\tau}_{A_{\mathcal{M}}}(i)$, $i = 1, \cdots, \hat{l}_n$, where $\hat{l}_n = \sum_{i=1}^{n} \mathbb{I}\{X_i \in S, \hat{Y}_i = 1\}$ is the total number of visits of the split chain to $A_{\mathcal{M}}$ up to time n. The approximated blocks are of the form:

$$\hat{\mathcal{B}}_0 = (X_1, \cdots, X_{\hat{\tau}_{A_{\mathcal{M}}}(1)}), \cdots, \hat{\mathcal{B}}_j = (X_{\hat{\tau}_{A_{\mathcal{M}}}(j)+1}, \cdots, X_{\hat{\tau}_{A_{\mathcal{M}}}(j+1)}), \cdots,$$

$$\hat{\mathcal{B}}_{\hat{l}_n - 1} = (X_{\hat{\tau}_{A_{\mathcal{M}}}(\hat{l}_n - 1)+1}, \cdots, X_{\hat{\tau}_{A_{\mathcal{M}}}(\hat{l}_n)}),\ \hat{\mathcal{B}}_{\hat{l}_n}^{(n)} = (X_{\hat{\tau}_{A_{\mathcal{M}}}(\hat{l}_n)+1}, \cdots, X_{n+1}).$$

Moreover, we denote by $\hat{n}_{A_{\mathcal{M}}} = \hat{\tau}_{A_{\mathcal{M}}}(\hat{l}_n) - \hat{\tau}_{A_{\mathcal{M}}}(1) = \sum_{i=1}^{\hat{l}_n - 1} l(\hat{B}_j)$ the total number of observations after the first and before the last pseudo-regeneration times. Let

$$\sigma_f^2 = \frac{1}{\mathbb{E}_{A_{\mathcal{M}}}(\tau_{A_{\mathcal{M}}})} \mathbb{E}_{A_{\mathcal{M}}} \left(\sum_{i=1}^{\tau_{A_{\mathcal{M}}}} \{f(X_i) - \mu(f)\}^2 \right)$$

be the asymptotic variance. Furthermore, we set that

$$\hat{\mu}_n(f) = \frac{1}{\hat{n}_{A_{\mathcal{M}}}} \sum_{i=1}^{\hat{l}_n - 1} f(\hat{B}_j), \quad \text{where } f(\hat{B}_j) = \sum_{i=1+\hat{\tau}_{A_{\mathcal{M}}}(j)}^{\hat{\tau}_{A_{\mathcal{M}}}(j+1)} f(X_i)$$

and

$$\hat{\sigma}_n^2(f) = \frac{1}{\hat{n}_{A_{\mathcal{M}}}} \sum_{i=1}^{\hat{l}_n - 1} \left\{ f(\hat{B}_i) - \hat{\mu}_n(f) l(\hat{B}_i) \right\}^2.$$

We briefly indicate that there exists a connection between α- mixing coefficients and regeneration times for Harris recurrent Markov chains. The strong α-mixing coefficient between σ-fields $\mathcal{A}$ and $\mathcal{B}$ is defined as

$$\alpha(\mathcal{A}, \mathcal{B}) := \sup_{(A,B) \in \mathcal{A} \times \mathcal{B}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|.$$

The strong mixing coefficients related to a sequence of random variables are defined by

$$\alpha(k) = \sup_n \sup_{A \in \hat{\xi}_n} \sup_{B \in \hat{\xi}^n} |\mathbb{P}_\mu(A \cap B) - \mathbb{P}_\mu(A)\mathbb{P}_\mu(B)|,$$

where $\hat{\xi}_n = \sigma(X_i, i \leq n)$ and $\hat{\xi}^n = \sigma(X_i, i \geq n)$. By Theorem 2 from Bolthausen (1982), we know that for stationary Harris chains if for some $\lambda \geq 0$ the sum

$\sum_m m^\lambda \alpha(m) < \infty$, then for all $B \in \mathcal{E}$ such that $\mu(B) > 0$ we have $\mathbb{E}_\mu(\tau_B^{1+\lambda}) < \infty$, where $\tau_B = \inf\{n \geq 1 : X_n \in B\}$. This result guarantees that the rate of decay of strong mixing coefficients is polynomial. This is a weaker condition, because usually the exponential rate of decay is assumed.

3. Bootstrap methods for Harris recurrent Markov chains

In this section we recall shortly some bootstrap methods for Harris recurrent Markov chains which are essential to establish our bootstrap versions of uniform central limit theorems for Markov chains. We formulate necessary assumptions which must be satisfied by the chain in order to our theory could work.

3.1. ARBB method

In this subsection we recall the Approximate block bootstrap algorithm (ARBB) introduced by Bertail and Clémençon (2006). The ARBB method allows to utilize the pseudo-regeneration structure of the split chain in order to generate the bootstrap blocks $B_1^*, \cdots, B_k^*$ which are obtained by resampling pseudo-regeneration data blocks $\hat{B}_1, \cdots, \hat{B}_{\hat{l}_n - 1}$. The algorithm allows to compute the estimate of the sample distribution of some statistic $T_n = T(\hat{B}_1, \cdots, \hat{B}_{\hat{l}_n - 1})$ with standarization $S_n = S(\hat{B}_1, \cdots, \hat{B}_{\hat{l}_n - 1})$. For the sake of clarity, we recall the ARBB bootstrap procedure below. The algorithm proceeds as follows:

Algorithm 3.1 (ARBB procedure)**.**

1. *Draw sequentially bootstrap data blocks $B_1^*, \cdots, B_k^*$ (we denote the length of the blocks by $l(B_j^*),\ j = 1, \cdots, k$) independently from the empirical distribution function*

$$\hat{\mathcal{L}}_n = \frac{1}{\hat{l}_n - 1} \sum_{i=1}^{\hat{l}_n - 1} \delta_{\hat{B}_i},$$

   where $\hat{B}_i,\ i = 1, \cdots, \hat{l}_n - 1$ are initial pseudo-regeneration blocks. We generate the bootstrap blocks until the joint length of the bootstrap blocks $l^*(k) = \sum_{i=1}^k l(B_i^*)$ exceeds n. We set $l_n^* = \inf\{k : l^*(k) > n\}$.
2. *Bind the bootstrap blocks from the step 1 and construct the ARBB bootstrap sample $X^{*(n)} = (X_1^*, \cdots, X_{l_n^*-1}^*)$.*
3. *Compute the ARBB statistic and its ARBB distribution, namely $T_n^* = T(X^{*(n)}) = T(B_1^*, \cdots, B_{l_n^*-1}^*)$ and its standarization $S_n^* = S(X^{*(n)}) = S(B_1^*, \cdots, B_{l_n^*-1}^*)$.*
4. *The ARBB distribution is given by*

$$H_{ARBB}(x) = \mathbb{P}^*(S_n^{*-1}(T_n^* - T_n) \leq x),$$

   where $\mathbb{P}^*$ is the conditional probability given the data.

We introduce few more pieces of notation. We denote by

$$n^*_{A_{\mathcal{M}}} = \sum_{i=1}^{l^*_n - 1} l(B^*_j)$$

the length of the bootstrap sample,

$$\mu^*_n(f) = \frac{1}{n^*_{A_{\mathcal{M}}}} \sum_{i=1}^{l^*_n-1} f(B^*_i) \text{ and } \sigma^{*2}_n(f) = \frac{1}{n^*_{A_{\mathcal{M}}}} \sum_{i=1}^{l^*_n-1} \{f(B^*_i) - \mu^*_n(f) l(B^*_j)\}^2.$$

3.2. Preliminary bootstrap results for Markov chains

Let (X_n) be a positive recurrent Harris Markov chain and $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative numbers that converges to zero. We impose the following assumptions on the chain (compare with Bertail and Clémençon (2006), page 700):

1. S is chosen so that $\inf_{x\in S} \phi(x) > 0$.
2. Transition density p is estimated by p_n at the rate α_n (usually we consider $\alpha_n = \frac{\log(n)}{n}$) for the mean squared error (MSE) when error is measured by the L^∞ loss over S^2.

Moreover, we assume the following conditions (for a comprehensive treatment on these assumptions the interested reader may refer to Bertail and Clémençon (2006)). Let $k \geq 2$ be a real number.
$\mathcal{H}_1(f, k, \nu)$. The small set S is such that

$$\sup_{x\in S} \mathbb{E}_x \left[\left(\sum_{i=1}^{\tau_S} |f(X_i)| \right)^k \right] < \infty$$

and

$$\mathbb{E}_\nu \left[\left(\sum_{i=1}^{\tau_S} |f(X_i)| \right)^k \right] < \infty.$$

$\mathcal{H}_2(k, \nu)$. The set S is such $\sup_{x\in S} \mathbb{E}_x(\tau_S^k) < \infty$ and $\mathbb{E}_\nu(\tau_S^k) < \infty$.
$\mathcal{H}_3$. $p(x, y)$ is estimated by $p_n(x, y)$ at the rate α_n for the MSE when error is measured by the L^∞ loss over $S \times S$:

$$\mathbb{E}_\nu \left(\sup_{(x,y)\in S\times S} |p_n(x, y) - p(x, y)|^2 \right) = O(\alpha_n), \quad \text{as } n \to \infty.$$

$\mathcal{H}_4$. The density ϕ is such that $\inf_{x\in S} \phi(x) > 0$.
$\mathcal{H}_5$. The transition density $p(x, y)$ and its estimate $p_n(x, y)$ are bounded by a constant $R < \infty$ over S^2.

Remark 3.1. *In the following, we assume that* $\alpha_n \sim (\frac{\log(n)}{n})^{s/s+1}$ *(see Bertail and Clémençon (2006) for more details).*

Before we establish main result of this paper we recall two theorems from Bertail and Clémençon (2006) that essentially establish the consistency of ARBB procedure for pseudo-regeneration blocks.

Theorem 3.2. *Assume that the conditions* [1] *and* [2] *are satisfied by the chain and* $\mathcal{H}_1(f,\rho,\nu)$, $\mathcal{H}_2(\rho,\nu)$ *with* $\rho \geq 4$, $\mathcal{H}_3, \mathcal{H}_4$ *and* $\mathcal{H}_5$ *hold. Then, as* $n \to \infty$ *we have*

$$\hat{\sigma}_n^2(f) \to \sigma_f^2 \quad \text{in } \mathbb{P}_\nu - \text{probability}$$

and

$$\hat{n}_{A_\mathcal{M}}^{1/2} \frac{\hat{\mu}_n(f) - \mu(f)}{\hat{\sigma}_n(f)} \to \mathcal{N}(0,1) \quad \text{in distribution under } \mathbb{P}_\nu.$$

Denote by $BL_1(\mathcal{F})$ the set of all 1-Lipschitz bounded functions on $l^\infty(\mathcal{F})$. We define the bounded Lipschitz metric on $l^\infty(\mathcal{F})$ as

$$d_{BL_1}(X,Y) = \sup_{b \in BL_1(l^\infty(\mathcal{F}))} |\mathbb{E}b(X) - \mathbb{E}b(Y)|; \quad X,Y \in l^\infty(\mathcal{F}).$$

Convergence in bounded Lipschitz metric is correspondent to weak convergence. Expectations of nonmeasurable elements are understood as outer expectations.

Definition 3.3. *We say that* $\mathbb{Z}_n^*$ *is weakly consistent if* $d_{BL_1}(\mathbb{Z}_n^*, \mathbb{Z}_n) \xrightarrow{P} 0$. *Analogously,* $\mathbb{Z}_n^*$ *is strongly consistent if* $d_{BL_1}(\mathbb{Z}_n^*, \mathbb{Z}_n) \xrightarrow{a.s.} 0$.

Theorem 3.4. *Under the hypotheses of the Theorem 3.2 , we have the following convergence in probability under* $\mathbb{P}_\nu$*:*

$$\Delta_n = \sup_{x \in \mathbb{R}} |H_{ARBB}(x) - H_\nu(x)| \to 0, \quad \text{as } n \to \infty,$$

where

$$H_\nu(x) = \mathbb{P}_\nu(x) \left(\hat{n}_{A_\mathcal{M}}^{1/2} \sigma_f^{-1} (\hat{\mu}_n(f) - \mu(f)) \leq x \right)$$

and

$$H_{ARBB}(x) = \mathbb{P}^* \left(n_{A_\mathcal{M}}^{*1/2} \hat{\sigma}_n^{-1}(f)(\mu_n^*(f) - \hat{\mu}_n(f)) \leq x | X^{(n+1)} \right).$$

In the following, the convergence $X_n \xrightarrow{P^*} X$ in $\mathbb{P}_{\nu-}$probability ($\mathbb{P}_\nu$ -a.s.) along the sample is understood as

$$\mathbb{P}^*(|X_n - X| > \epsilon | X^{(n+1)}) \xrightarrow{n \to \infty} 0 \quad \text{in } \mathbb{P}_\nu - \text{probability} \ \ (\mathbb{P}_\nu \text{ -a.s.}).$$

4. Uniform bootstrap central limit theorems for Markov chains

To establish the uniform bootstrap CLT over permissible, uniformly bounded classes of functions $\mathcal{F}$, we need to be sure that the size of $\mathcal{F}$ is not too large (it is typical requirement when considering uniform asymptotic results for empirical processes indexed by classes of functions). To control the size of $\mathcal{F}$, we require the finiteness of its *covering number* $N_p(\epsilon, Q, \mathcal{F})$ which is interpreted as the minimal number of balls with radius ϵ needed to cover $\mathcal{F}$ in the norm $L_p(Q)$ and Q is a measure on E with finite support. Moreover, we impose the finiteness of the *uniform entropy integral* of $\mathcal{F}$, namely

$$\int_0^\infty \sqrt{\log N_2(\epsilon, \mathcal{F})}d\epsilon < \infty, \quad \text{where } \; N_2(\epsilon, \mathcal{F}) = \sup_Q N_2(\epsilon, Q, \mathcal{F}).$$

For the sake of completeness we recall below Theorem 5.9 from Levental (1988) which is crucial to establish uniform bootstrap CLT in general Harris case.

Theorem 4.1. *Let (X_n) be a positive recurrent Harris chain taking values in $(E, \mathcal{E})$. Let μ be the invariant probability measure for (X_n). Assume further that $\mathcal{F}$ is a uniformly bounded class of measurable functions on E and*

$$\int_0^\infty \sqrt{\log N_2(\epsilon, \mathcal{F})}d\epsilon < \infty.$$

If $\sup_{x\in A} \mathbb{E}_x(\tau_A)^{2+\gamma} < \infty$ ($\gamma > 0$ fixed), where A is atomic set for the chain, then the empirical process $Z_n(f) = n^{1/2}(\mu_n - \mu)(f)$ converges weakly as a random element of $l^\infty(\mathcal{F})$ to a gaussian process G indexed by $\mathcal{F}$ whose sample paths are bounded and uniformly continuous with respect to the metric $L_2(\mu)$.

4.1. Main asymptotic results

In this subsection we establish the bootstrap uniform central limit theorem over permissible, uniformly bounded classes of functions which satisfy the uniform entropy condition.

Theorem 4.2. *Suppose that (X_n) is positive recurrent Harris Markov chain and the assumptions from the Theorem 3.4 are satisfied by (X_n). Assume further that $\mathcal{F}$ is a permissible, uniformly bounded class of functions and the following uniformity condition holds*

$$\int_0^\infty \sqrt{\log N_2(\epsilon, \mathcal{F})}d\epsilon < \infty. \tag{4.1}$$

Then the process

$$\mathbb{Z}_n^* = n_{A_\mathcal{M}}^{*1/2} \left[\frac{1}{n^*_{A_\mathcal{M}}} \sum_{i=1}^{l_n^*-1} f(B_i^*) - \frac{1}{\hat{n}_{A_\mathcal{M}}} \sum_{i=1}^{\hat{l}_n-1} f(\hat{B}_i) \right] \tag{4.2}$$

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

Proof. The proof is based on the bootstrap central limit theorem introduced by Giné and Zinn (1990). To prove the weak convergence of the process $\mathbb{Z}_n^*$ we need to show

1. Finite dimensional convergence of distributions of $\mathbb{Z}_n^*$ to G.
2. Stochastic asymptotic equicontinuity in probability under $\mathbb{P}_\nu$ with respect to the totally bounded semimetric ρ on $\mathcal{F}$.

Firstly, we prove that $(\mathbb{Z}_n^*(f_{i1}), \cdots, \mathbb{Z}_n^*(f_{ik}))$ converges weakly in probability to $(G(f_{i1}), \cdots, G(f_{ik}))$ for every fixed finite collection of functions $\{f_{i1}, \cdots, f_{ik}\} \subset \mathcal{F}$. Denote by $\xrightarrow{L}$ the weak convergence in law in the sense of Hoffmann-Jørgensen. We want to show that for any fixed collection $(a_1, \cdots, a_k) \in \mathbb{R}$ we have

$$\sum_{j=1}^{k} a_j \mathbb{Z}_n^*(f_{ij}) \xrightarrow{L} \mathcal{N}(0, \gamma^2) \quad \text{in probability under } \mathbb{P}_\nu,$$

where

$$\gamma^2 = \sum_{j=1}^{k} a_j^2 Var(\mathbb{Z}_n(f_{ij})) + \sum_{s \neq r} a_i a_j Cov(\mathbb{Z}_n(f_{is}), \mathbb{Z}_n(f_{ir})).$$

Let $h = \sum_{j=1}^{k} a_j f_{ij}$. By linearity of h and Theorem 4.1 we conclude that

$$\mathbb{Z}_n(h) \xrightarrow{L} G(h). \tag{4.3}$$

The above convergence of $\mathbb{Z}_n(h)$ coupled with the Theorems 3.2 and 3.4 guarantee that $\mathbb{Z}_n^*(h) \xrightarrow{L} G(h)$ in probability under $\mathbb{P}_\nu$. Thus, the finite dimensional convergence for the $\mathbb{Z}_n^*(f),\ f \in \mathcal{F}$ is established.

To verify [2] we need to check if for every $\epsilon > 0$

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}^*(\|\mathbb{Z}_n^*\|_{\mathcal{F}_\delta} > \epsilon) = 0 \ \text{ in probability under } \mathbb{P}_\nu, \tag{4.4}$$

where $\|R\|_{\mathcal{F}_\delta} := \sup\{|R(f) - R(g)| : \rho(f, g) < \delta\}$ and $R \in l^\infty(\mathcal{F})$. Moreover, $\mathcal{F}$ must be totally bounded in $L_2(\mu)$. In fact, the latter was shown by Levental (1988). For the reader's convenience we repeat the reasoning from the mentioned paper.

Consider class of functions $\mathcal{H} = \{|f - g| : f, g \in \mathcal{F}\}$. Denote by Q_n the n-th empirical measure of an i.i.d. process whose law is μ. Using basic properties of covering numbers we obtain that $N_1(\epsilon, \mathcal{G}, Q_n) \leq \left(N_2\left(\frac{\epsilon}{2}, \mathcal{F}\right)\right)^2 < \infty$ and thus by the SLLN for Q_n (see Theorem 3.6 in Levental (1988)) we have that

$$\sup_{h \in \mathcal{H}} |(Q_n - \mu)(h)| \to 0 \ \text{ a.s.}(\mu).$$

Since $\mathcal{F}$ is totally bounded in $L_1(Q)$ for every measure Q with finite support it follows that is totally bounded in $L_1(\mu)$. Moreover, one can show that if an envelope of $\mathcal{F}$ is in $L_2(\mu)$, then $\mathcal{F}$ is totally bounded in $L_2(\mu)$.

In order to show (4.4), firstly, we replace the random numbers $n^*_{A_{\mathcal{M}}}$ and l^*_n by their deterministic equivalents. By the same arguments as in the proof of the Theorem 3.4 (see Bertail and Clémençon (2006), page 710 for details) we have the following convergences

$$\frac{l(B_j^*)}{n} \xrightarrow{P^*} 0 \text{ and } \frac{n^*_{A_{\mathcal{M}}}}{n} \xrightarrow{P^*} 1$$

in $\mathbb{P}_\nu-$probability along the sample path as $n \to \infty$ and

$$\frac{l_n^*}{n} - \mathbb{E}_{A_{\mathcal{M}}}(\tau_{A_{\mathcal{M}}})^{-1} \xrightarrow{P^*} 0$$

in $\mathbb{P}_\nu-$probability along the sample path as $n \to \infty$. Thus, we conclude that

$$\begin{aligned}
\mathbb{Z}_n^*(f) &= \sqrt{n^*_{A_{\mathcal{M}}}}\left[\frac{1}{n^*_{A_{\mathcal{M}}}}\sum_{i=1}^{l_n^*-1} f(B_i^*) - \frac{1}{\hat{n}_{A_{\mathcal{M}}}}\sum_{i=1}^{\hat{l}_n-1} f(\hat{B}_i)\right] \\
&= \frac{1}{\sqrt{n^*_{A_{\mathcal{M}}}}}\left[\sum_{i=1}^{l_n^*-1}\{f(B_i^*) - \hat{\mu}_n(f)l(B_i^*)\}\right] \\
&= \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{1+\left\lfloor \frac{n}{\mathbb{E}_{A_{\mathcal{M}}}(\tau_A)}\right\rfloor}\{f(B_i^*) - \hat{\mu}_n(f)l(B_i^*)\}\right] + o_{\mathbb{P}^*}(1),
\end{aligned}$$

where $\lfloor x \rfloor$ is an integer part of $x \in \mathbb{R}$. The preceding reasoning allows us to switch to the process

$$\mathbb{U}_n^*(f) = \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{1+\left\lfloor \frac{n}{\mathbb{E}_{A_{\mathcal{M}}}(\tau_A)}\right\rfloor}\{f(B_i^*) - \hat{\mu}_n(f)l(B_i^*)\}\right].$$

Observe, that $\{f(B_i^*) - \hat{\mu}_n(f)l(B_i^*)\}_{i\geq 1}$ forms the sequence of i.i.d. random variables.

Next, take $h = f - g$. Denote by $w(n) = 1 + \left\lfloor \frac{n}{\mathbb{E}_{A_{\mathcal{M}}}(\tau_A)}\right\rfloor$ and $Y_i = l(B_i^*) - \hat{\mu}_n(f)l(B_i^*)$. We have the following inequality (by the fact that Y_i's are i.i.d.)

$$\mathbb{P}^*(\|\mathbb{U}_n^*(h)\|_{\mathcal{F}_\delta} > \epsilon) \leq w(n)\mathbb{P}^*\left(\frac{1}{\sqrt{n}}\|h(B_1^*) - l(B_1^*)\hat{\mu}_{n,h}\|_{\mathcal{F}_\delta} > \epsilon\right).$$

The right hand side of the above inequality is bounded by

$$w(n)\mathbb{P}^*\left(\|h(B_1^*)\|_{\mathcal{F}_\delta} > \frac{\sqrt{n}\epsilon}{2}\right) + w(n)\mathbb{P}^*\left(|l(B_1^*)|\|\hat{\mu}_{n,h}\|_{\mathcal{F}_\delta} > \frac{\sqrt{n}\epsilon}{2}\right) = I + II.$$

In the following, we investigate the asymptotic behaviour of I and II. Some of the reasoning relies on the useful proposition from Radulović (2004).

Proposition 4.3. *For any random variable W, such that $\mathbb{E}W^2 < \infty$, there exists a positive increasing function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ such that*

$$\lim_{x\to\infty} \frac{\phi(x)}{x^2} = +\infty \ \textit{ and } \ \mathbb{E}\phi(W) < \infty.$$

Remark 4.1. *The sketch of the proof of the above Proposition is moved to the Appendix section.*

By Markov's inequality, we have that

$$w(n)\mathbb{P}^*\left(\|h(B_1^*)\|_{\mathcal{F}_\delta} > \frac{\sqrt{n}\epsilon}{2}\right) \le w(n)\frac{\mathbb{E}^*(\phi(\|h(B_1^*)\|_{\mathcal{F}_\delta}))}{\phi(\frac{\sqrt{n}\epsilon}{2})}$$

By the Proposition 4.3 we conclude that

$$\frac{w(n)}{\phi\left(\sqrt{n}/2\right)} = \frac{w(n)}{n} \cdot \frac{n}{\phi\left(\sqrt{n}/2\right)} \to 0 \ \text{ a.s.}$$

since $\frac{w(n)}{n} \le 1$. Note also that

$$\mathbb{E}^*(\phi(\|h(B_1^*)\|_{\mathcal{F}_\delta})) \le \mathbb{E}^*(\phi(|2F(B_1^*)|)) < \infty \ \text{ a.s.}$$

since $\mathcal{F}$ is uniformly bounded. Thus,

$$w(n)\mathbb{P}^*\left(\|h(B_1^*)\|_{\mathcal{F}_\delta} > \frac{\sqrt{n}\epsilon}{2}\right) \to 0 \text{ a.s.}$$

Next, we investigate the asymptotic behaviour of II. By Markov's inequality, we have

$$w(n)\mathbb{P}^*\left(|l(B_1^*)|\|\hat{\mu}_{n,h}\|_{\mathcal{F}_\delta} > \frac{\sqrt{n}\epsilon}{2}\right) \le 4w(n)\frac{\mathbb{E}^*(|l(B_1^*)|)^2\|\hat{\mu}_{n,h}\|^2_{\mathcal{F}_\delta}}{n}.$$

We know that $\frac{w(n)}{n} \le 1$ and $\|\hat{\mu}_{n,h}\|_{\mathcal{F}_\delta} \to 0$ in $\mathbb{P}_\nu-$probability because of the stochastic equicontinuity of the original process $\mathbb{Z}_n$. Moreover, it is proven in Bertail and Clémençon (2006) that

$$\mathbb{E}^*\left(l(B_1^*)^2|X^{(n+1)}\right) \to \mathbb{E}_{A_\mathcal{M}}(\tau^2_{A_\mathcal{M}}) < \infty$$

in $\mathbb{P}_\nu-$probability along the sample as $n \to \infty$. Thus,

$$w(n)\mathbb{P}^*\left(|l(B_1^*)|\|\hat{\mu}_{n,h}\|_{\mathcal{F}_\delta} > \frac{\sqrt{n}\epsilon}{2}\right) \to 0$$

in $\mathbb{P}_\nu-$probability along the sample as $n \to \infty$.

The above reasoning implies that (4.4) holds. We have checked that both conditions [1] and [2] are satisfied by $\mathbb{Z}_n^*$. Thus, we can apply the bootstrap CLT proposed by Giné and Zinn (1990) which yields the desired result. □

Remark 4.2. *Theorem 4.2 is a generalization of the Theorem 2.2 from Radulović (2004) for countable Markov chains. Note that the reasoning from the proof of the above theorem can be directly applied to the proof of Radulović's result. The part concerning the proof of the asymptotic stochastic equicontinuity of the bootstrap version of the empirical process indexed by uniformly bounded class of functions $\mathcal{F}$ can be significantly simplified. As shown in the proof of the Theorem 4.2, we can switch from the process $\mathbb{Z}_n^*(f)_{f\in\mathcal{F}} := \sqrt{n^*}\{\mu_{n^*}(f) - \mu_{n_A}(f)\}$, where $n_A = \tau_A(l_n) - \tau_A$ to the process*

$$\mathbb{U}_n^*(f) = \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{1+\left\lfloor \frac{n}{\mathbb{E}_A(\tau_A)}\right\rfloor}\{f(B_i^*) - \mu_{n_A}(f)l(B_i^*)\}\right]$$

and the standard probability inequalities applied to the i.i.d. blocks of data yield the result.

In the following, we show that we can weaken the assumption of uniform boundedness imposed on the class $\mathcal{F}$. By the results of Tsai (1998), it is sufficient that $\mathcal{F}$ has an envelope in $L_2(\mu)$, then the uniform bootstrap central limit theorem holds.

Theorem 4.4. *Suppose that (X_n) is positive recurrent Harris Markov chain and the assumptions from the Theorem 3.4 are satisfied by (X_n). Assume further that $\mathcal{F}$ is a permissible class of functions and such that the envelope F satisfies*

$$\mathbb{E}_{A_\mathcal{M}}\left[\sum_{\tau_{A_\mathcal{M}}<j\le\tau_{A_\mathcal{M}}(2)} F(X_j)\right]^{2+\gamma} < \infty, \quad \gamma > 0 \text{ (fixed)}. \tag{4.5}$$

Suppose, that the following uniformity condition holds

$$\int_0^\infty \sqrt{\log N_2(\epsilon,\mathcal{F})}\,d\epsilon < \infty. \tag{4.6}$$

Then the process

$$\mathbb{Z}_n^* = n_{A_\mathcal{M}}^{*1/2}\left[\frac{1}{n^*_{A_\mathcal{M}}}\sum_{i=1}^{l_n^*-1} f(B_i^*) - \frac{1}{\hat{n}_{A_\mathcal{M}}}\sum_{i=1}^{\hat{l}_n-1} f(\hat{B}_i)\right] \tag{4.7}$$

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

Proof. The proof of Theorem 4.4 goes analogously to the proof of the Theorem 4.2 with few natural modifications. We indicate the critical points where the changes are necessary. The notation remains in the agreement with the previous theorem.

- Theorem 4.3 from Tsai (1998) establishes the weak convergence

$$\mathbb{Z}_n(h) \xrightarrow{L} G(h).$$

- According to Bertail and Clémençon (2006), Theorem 3.4 is also true when f is unbounded (see Bertail and Clémençon (2006), page 706 for details). Thus, the finite dimensional convergence of distributions of $\mathbb{Z}_n^*$ to the right gaussian process is ensured.
- It is shown in Tsai (1988) that $\mathcal{F}$ is totally bounded in $L_2(\mu)$ when $\mathcal{F}$ fulfills only the condition that the envelope F is in $L_2(\mu)$ (see Tsai (1998), page 9 for details).
- The finiteness of the $\mathbb{E}^*(\phi(|2F(B_1^*)|))$ in $\mathbb{P}_v-$probability along the sample as $n \to \infty$ follows from the Proposition 4.3. We know that if the condition (4.5) on the envelope F holds, there exists a positive increasing function $\phi:\ \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$\phi(x) = x^{2+\gamma} \quad \text{and} \quad \mathbb{E}_{A_{\mathcal{M}}}\left[\sum_{\tau_{A_{\mathcal{M}}} < j \le \tau_{A_{\mathcal{M}}}(2)} \phi(F(X_j))\right]^{2+\gamma} < \infty.$$

□

Remark 4.3. *It is noteworthy that the uniform central limit theorem for Harris recurrent Markov chains in Tsai (1998) holds with the weaker condition on the envelope F, i.e.*

$$\mathbb{E}_{A_{\mathcal{M}}}\left[\sum_{\tau_{A_{\mathcal{M}}} < j \le \tau_{A_{\mathcal{M}}}(2)} F(X_j)\right]^{2} < \infty.$$

However, in the unbounded version of the bootstrap uniform central limit theorem for Harris Markov chains we need to assume the finiteness of the $(2+\gamma)$-th moment in order to show the finiteness of the $\mathbb{E}^*(\phi(|2F(B_1^*)|))$ in $\mathbb{P}_v-$probability along the sample as $n \to \infty$.

5. Bootstrapping Fréchet differentiable functionals

Robust statistics provides tools to deal with data when we suspect that they include a small proportion of outliers. Robust statistical methods are applied to the solution of many problems such as estimation of regression parametres, estimation of scale and location. One of the key concepts of robust statistics when detecting the outliers in the data is an influence function. In the i.i.d. setting, the influence function measures the change in the value of some functional $\phi(P)$ if we replace some infinitesimally small part of P by a pointmass x (see van der Vaart (2000) for a detailed treatment of these issues in the i.i.d. framework). Generalizing the concepts of robustness and influence function into dependent case is a challeging task (see Bertail and Clémençon (2015) and references therein). In a Markovian setting, one can measure the influence of (approximate) regeneration data blocks instead of single observations. The regenerative approach proposed by Bertail and Clémençon (2015) naturally leads to central limit and convolution theorems.

In our framework, we show how we can use results from the previous section to yield the bootstrap uniform central limit theorems for general differentiable functionals over uniformly bounded classes (and with an envelope in $L_2(\mu)$) of functions $\mathcal{F}$.

5.1. Preliminary assumptions and remarks

In robust statistics the influence function plays a crucial role to detect outliers in data. Functions and estimators which have an unbounded influence function should be carefully investigated, because the small proportion of the observations would have too much influence on the estimator.

Let's make our considerations rigorous. We denote by $\mathcal{P}$ the set of all probability measures on E. We keep the notation in agreement with notation introduced in Bertail and Clémençon (2015).

The classical definition of the influence function is provided below.

Definition 5.1. *Let $(\vartheta, \|\cdot\|)$ be a separable Banach space. Let $T : \mathcal{P} \to \vartheta$ be a functional on $\mathcal{P}$. If the limit*

$$\frac{T((1-t)\mu + t\delta_x) - T(\mu)}{t}, \quad \text{as } t \to 0$$

is finite for all $\mu \in \mathcal{P}$ and for any $x \in E$, then we say that the influence function $T^{(1)} : \mathcal{P} \to \vartheta$ of the functional T is well-defined and for all $x \in E$

$$T^{(1)}(x,\mu) = \lim_{t\to 0} \frac{T((1-t)\mu + t\delta_x) - T(\mu)}{t}.$$

In the following, we recall the definition of Fréchet derivative which is an important concept in robust statistics. In particular, Fréchet differentiability ensures the existence of the influence function. Let d be some metric on $\mathcal{P}$.

Definition 5.2. *We say that the functional $T : \mathcal{P} \to \mathbb{R}$ is Fréchet differentiable at $\mu_0 \in \mathcal{P}$ for a metric d, if there exists a continuous linear operator DT_{μ_0} (from the set of signed measures of the form $\mu - \mu_0$ in $(\vartheta, \|\cdot\|)$) and a function $\epsilon^{(1)}(\cdot, \mu_0) : \mathbb{R} \to (\vartheta, \|\cdot\|)$, which is continuous at 0 and $\epsilon^{(1)}(0, \mu_0) = 0$ such that*

$$\forall\ \mu \in \mathcal{P},\ \ T(\mu) - T(\mu_0) = DT_{\mu_0}(\mu - \mu_0) + R^{(1)}(\mu, \mu_0),$$

where $R^{(1)}(\mu, \mu_0) = d(\mu, \mu_0)\epsilon^{(1)}(d(\mu, \mu_0), \mu_0)$. Furthermore, we say that T has an influence function $T^{(1)}(\cdot, \mu_0)$ if the following representation holds for DT_{μ_0} :

$$\forall\ \mu_0 \in \mathcal{P}, \quad DT_{\mu_0}(\mu - \mu_0) = \int_E T^{(1)}(x, \mu_0)\mu(dx).$$

In the context of empirical processes indexed by classes of functions, when one want to derive the uniform central limit theorems for generally differentiable functionals the appropriate choice of metric is the crucial point. We need to choose the metric carefully in order to precisely control the distance $d(\mu_n, \mu)$ and the remainder $R^{(1)}(\mu_n, \mu)$. In our framework we have decided to work with a generalization of the Kolmogorov's distance which is defined as follows:

Definition 5.3. *Let $\mathcal{H}$ be a class of real-valued functions (we do not impose the measurability condition as one can work with outer measures and the Hoffmann-Jørgensen (1991) convergence). We define a distance*

$$d_{\mathcal{H}}(P, Q) := \sup_{h \in \mathcal{H}} \left| \int h d(P - Q) \right| \tag{5.1}$$

for any $P, Q \in \mathcal{P}$.

The choice of metric defined in (5.1) is inspired by the arguments given by Barbe and Bertail (1995) and Dudley (1990). Essentially, one may want to work with metric $d_{\mathcal{H}}$ because it enables very precise control of the distance $d(\mu_n, \mu)$. Moreover, in many cases we can find a class of functions $\mathcal{H}$, which makes the functionals Fréchet diferrentiable for $d_{\mathcal{H}}$. The latter is a significant advantage since choice of metric that guarantees Fréchet differentiability of functionals is usually challenging (see Barbe and Bertail (1995) and Dudley (1990) for an extensive treatment on this subject).

Note, that permissible, uniformly bounded (or with an envelope in $L_2(\mu)$) classes of functions $\mathcal{F}$ fulfill the conditions imposed on the class $\mathcal{H}$. Thus, we can ease the notation and write $d_{\mathcal{F}}$ for the distance defined by (5.1).

5.2. Bootstrap uniform central limit theorems for Fréchet differentiable functionals

In this subsection, we show how the results from Levental (1988), Tsai (1998) and from the previous section yield the uniform bootstrap central limit theorems for Fréchet differentiable functionals. Before we formulate the theorems, we briefly recall the notation. In general Harris case,

$$\mu_n^* = \frac{1}{n^*_{A_{\mathcal{M}}}} \sum_{i=1}^{l_n^*-1} f(B_i^*) \text{ and } \hat{\mu}_n = \frac{1}{\hat{n}_{A_{\mathcal{M}}}} \sum_{i=1}^{\hat{l}_n-1} f(\hat{B}_i),$$

where $\hat{B}_i,\ i = 1, \cdots, \hat{l}_n - 1$ are pseudo-regeneration blocks. In regenerative case, the empirical mean is of the form

$$\mu_n = \frac{1}{n_A} \sum_{i=1}^{l_n-1} f(B_i).$$

The crucial observation in order to establish the results is, that as long as we can control the distance $d_{\mathcal{F}}(\mu_n^*, \hat{\mu}_n)$ (we require it would be sufficiently small), we can control the remainder term $R^{(1)}(\mu_n^*, \hat{\mu}_n)$. By the uniform central limit theorem, the linear part of the $T(\mu_n^*) - T(\hat{\mu}_n)$ is converging weakly to a desired gaussian process which yields our result.

Theorem 5.4. *Let $\mathcal{F}$ be a permissible, uniformly bounded class of functions, such that*

$$\int_0^\infty \sqrt{\log N_2(\epsilon, \mathcal{F})} d\epsilon < \infty.$$

Suppose that the conditions of Theorem 4.2 hold and $T : \mathcal{P} \to \mathbb{R}$ is Fréchet differentiable functional at μ. Then, in general Harris positive recurrent case, we have that $n^{1/2}(T(\mu_n^*) - T(\hat{\mu}_n))$ converges weakly in $l^\infty(\mathcal{F})$ to a gaussian process G_μ indexed by $\mathcal{F}$, whose sample paths are bounded and uniformly continuous with respect to the metric $L_2(\mu)$.

Remark 5.1. *It is obvious that the above theorem works also in the regenerative case. Replace $\mathcal{A}_\mathcal{M}$ and the $\hat{\mu}_n$ for the split chain by $\mathcal{A}$ and μ_n respectively. Then, under the assumptions from Theorem 5.4, we have the weak convergence in $l^\infty(\mathcal{F})$ to the gaussian process indexed by $\mathcal{F}$, whose sample paths are bounded and uniformly continuous with respect to the metric $L_2(\mu)$.*

Proof. Without loss of generality, we assume that $\mathbb{E}_\mu T^{(1)}(x, \mu) = 0$. By the Fréchet differentiability formulated in definition 5.2 we have

$$T(\hat{\mu}_n) - T(\mu) = DT_\mu(\hat{\mu}_n - \mu) + d_\mathcal{F}(\hat{\mu}_n, \mu)\epsilon^{(1)}(d_\mathcal{F}(\hat{\mu}_n, \mu), \mu) \tag{5.2}$$

and

$$T(\mu_n^*) - T(\mu) = DT_\mu(\mu_n^* - \mu) + d_\mathcal{F}(\mu_n^*, \mu)\epsilon^{(1)}(d_\mathcal{F}(\mu_n^*, \mu), \mu). \tag{5.3}$$

Thus,

$$\sqrt{n}(T(\mu_n^*) - T(\hat{\mu}_n)) = \sqrt{n}\left(DT_{\hat{\mu}_n}(\mu_n^* - \hat{\mu}_n)\right) + \sqrt{n}\left(d_\mathcal{F}(\hat{\mu}_n, \mu)\epsilon^{(1)}(d_\mathcal{F}(\hat{\mu}_n, \mu), \mu)\right)$$
$$+ \sqrt{n}\left(d_\mathcal{F}(\mu_n^*, \mu)\epsilon^{(1)}(d_\mathcal{F}(\mu_n^*, \mu), \mu)\right).$$

We show that $d_\mathcal{F}(\hat{\mu}_n, \mu)$ and $d_\mathcal{F}(\mu_n^*, \mu)$ are of order $O_{\mathbb{P}_\nu}(n^{-1/2})$. Theorem 4.1 guarantees that

$$\sqrt{n} d_\mathcal{F}(\hat{\mu}_n, \mu) \xrightarrow{L} \sup_{f \in \mathcal{F}} |G(f)|, \quad \text{as } n \to \infty,$$

where G is gaussian process whose sample paths are bounded and uniformly continuous with respect to the metric $L_2(\mu)$. Thus, $d_\mathcal{F}(\hat{\mu}_n, \mu) = O_{\mathbb{P}_\nu}(n^{-1/2})$.

Next, observe that

$$d_\mathcal{F}(\mu_n^*, \mu) \le d_\mathcal{F}(\mu_n^*, \hat{\mu}_n) + d_\mathcal{F}(\hat{\mu}_n, \mu).$$

From the Theorem 4.2 we conclude that

$$\sqrt{n} d_\mathcal{F}(\mu_n^*, \hat{\mu}_n) \xrightarrow{L^*} \sup_{f \in \mathcal{F}} |G(f)|, \quad \text{as } n \to \infty.$$

Thus, $d_\mathcal{F}(\mu_n^*, \hat{\mu}_n) = O_{P^*}\left(n^{-1/2}\right)$.

Remark 5.2. *Note that G and G_μ are not the same gaussian processes!*

We show that $d_{\mathcal{F}}(\mu_n^*, \hat{\mu}_n) = O_{\mathbb{P}_\nu}\left(n^{-1/2}\right)$. Indeed, consider the sequence S_n of order $O_{P^*}(1)$ in $\mathbb{P}_\nu-$probability along the sample, i.e.

$$\lim_{T\to\infty} \limsup_{n\to\infty} \mathbb{P}^*\{|S_n| \geq T\} \to 0 \quad \text{in } \mathbb{P}_\nu - \text{probability along the sample.}$$

Then,

$$\begin{aligned}\lim_{T\to\infty} \limsup_{n\to\infty} \mathbb{P}_\nu\{|S_n| \geq T\} &= \lim_{T\to\infty} \limsup_{n\to\infty} \mathbb{E}_\nu\left[\mathbb{P}^*\{|S_n| \geq T\}\right] \\ &\leq \lim_{T\to\infty} \mathbb{E}_\nu\left[\limsup_{n\to\infty} \mathbb{P}^*\{|S_n| \geq T\}\right] \\ &= \mathbb{E}_\nu\left[\lim_{T\to\infty} \limsup_{n\to\infty} \mathbb{P}^*\{|S_n| \geq T\}\right] = 0\end{aligned}$$

by the dominated convergence theorem and the Fatou's lemma. Thus, $d_{\mathcal{F}}(\mu_n^*, \hat{\mu}_n) = O_{\mathbb{P}_\nu}(n^{-1/2})$ and $d_{\mathcal{F}}(\mu_n^*, \mu) = O_{\mathbb{P}_\nu}(n^{-1/2})$.

Next, we scale (5.2) by $\sqrt{n}$:

$$\sqrt{n}(T(\hat{\mu}_n) - T(\mu)) = \sqrt{n}(DT_\mu(\hat{\mu}_n - \mu)) + o_{\mathbb{P}_\nu}(1)$$

and apply Theorem 4.2. Observe that the linear part in the above equation is gaussian as long as $0 < \mathbb{E}_\mu T^{(1)}(X_i, \mu)^2 \leq C_1^2(\mu)\mathbb{E}_\mu F^2(X) < \infty$ (see Barbe and Bertail (1995), chapter I for details), but that assumption is of course fulfilled since $\mathcal{F}$ is uniformly bounded. Thus, the following weak convergence in $l^\infty(\mathcal{F})$ holds:

$$\begin{aligned}\sqrt{n}(T(\hat{\mu}_n) - T(\mu)) &= \sqrt{n}(DT_\mu(\hat{\mu}_n - \mu)) + o_{\mathbb{P}_\nu}(1) \\ &= \sqrt{n}\int_E T^{(1)}(x,\mu)(\hat{\mu}_n - \mu)d(x) \\ &= \sqrt{n}\left[\frac{1}{\hat{n}_{\mathcal{A}_\mathcal{M}}}\sum_{i=1}^{\hat{n}_{\mathcal{A}_\mathcal{M}}} T^{(1)}(X_i,\mu) - 0\right] + o_{\mathbb{P}_\nu}(1) \xrightarrow{L} DT_\mu G_\mu.\end{aligned}$$

By the previous discussion, we also have

$$\begin{aligned}\sqrt{n}(T(\mu_n^* - T(\mu))) &= \sqrt{n}(DT_\mu(T(\mu_n^* - \mu)) + o_{\mathbb{P}_\nu}(1) \\ &= \sqrt{n}\int_E T^{(1)}(x,\mu)(\mu_n^* - \mu)d(x) \\ &= \sqrt{n}\left[\frac{1}{n^*_{\mathcal{A}_\mathcal{M}}}\sum_{i=1}^{n^*_{\mathcal{A}_\mathcal{M}}} T^{(1)}(X_i^*,\mu) - 0\right] + o_{\mathbb{P}_\nu}(1).\end{aligned}$$

The above convergences yield

$$\begin{aligned}\sqrt{n}[T(\mu_n^*) - T(\hat{\mu}_n)] &= \sqrt{n}\left[\frac{1}{n^*_{\mathcal{A}_\mathcal{M}}}\sum_{i=1}^{n^*_{\mathcal{A}_\mathcal{M}}} T^{(1)}(x_i^*,\mu) - \frac{1}{\hat{n}_{\mathcal{A}_\mathcal{M}}}\sum_{i=1}^{\hat{n}_{\mathcal{A}_\mathcal{M}}} T^{(1)}(x_i,\mu)\right] + o_{\mathbb{P}_\nu}(1) \\ &\xrightarrow{L} DT_\mu G_\mu\end{aligned}$$

and this completes the proof. □

Theorem 5.4 can be easily generalized to the case when $\mathcal{F}$ is unbounded and has the envelope in $L_2(\mu)$.

Theorem 5.5. *Let $\mathcal{F}$ be a permissible class of functions such that the envelope F satisfies*

$$\mathbb{E}_{A_{\mathcal{M}}}\left[\sum_{\tau_{A_{\mathcal{M}}}<j\leq\tau_{A_{\mathcal{M}}}(2)} F(X_j)\right]^{2+\gamma} < \infty, \;\; \gamma > 0 \text{ (fixed)}. \tag{5.4}$$

Suppose, that the following uniformity condition holds

$$\int_0^\infty \sqrt{\log N_2(\epsilon, \mathcal{F})}d\epsilon < \infty. \tag{5.5}$$

Assume further that the conditions of Theorem 4.4 hold and that $T : \mathcal{P} \to \mathbb{R}$ is Fréchet differentiable functional at μ. Then, in general Harris positive recurrent case, we have that $n^{1/2}(T(\mu_n^*) - T(\hat{\mu}_n))$ converges weakly in $l^\infty(\mathcal{F})$ to a gaussian process G_μ indexed by $\mathcal{F}$, whose sample paths are bounded and uniformly continuous with respect to the metric $L_2(\mu)$.

The proof of Theorem 5.5 follows analogously to the proof of Theorem 5.4. Apply the results of Tsai (1998) and Theorem 4.4 instead of Levental's (1988) and Theorem 4.2 to control the remainder terms. Then, the reasoning goes line by line as in the proof of Theorem 5.4.

Remark 5.3. *In particular, Theorem 5.5 is also true in the regenerative case. Replace $\hat{\mu}_n$ and $A_{\mathcal{M}}$ by μ_n and A. The proof goes analogously as in the preceding theorems.*

6. Conclusion

In this paper, we have shown how the regenerative properties of Markov chains can generalize some concepts in nonparametric statistics from i.i.d. to dependent case. We have shown that uniform bootstrap functional central limit theorem holds over permissible, uniformly bounded classes of functions. We have proved that the uniform boundedness assumption imposed on $\mathcal{F}$ can be weakened and it is feasible to require that $\mathcal{F}$ has an envelope in $L_2(\mu)$. We have worked with Markov chains on the general state space, but our results can be directly applied to Markov chains on countable state space. Thus, some proofs of the already existing results for the countable case, can be simplified when just applying the methodology introduced in this paper.

The bootstrap asymptotic results for empirical processes indexed by $\mathcal{F}$ naturally lead to bootstrap central limit theorems for Fréchet differentiable functionals. We have shown that bootstrap uniform CLTs hold in the bounded and the unbounded case over $\mathcal{F}$. Similar approach can be also applied to Hadamard differentiable functionals in order to establish analogous asymptotic results to presented in this paper.

Acknowledgement

I would like to thank my advisor, Patrice Bertail, for insightful remarks, inspiring discussions and guidance when I was working on this paper.

Appendix

In the small Appendix section we give the short proof of the Proposition 4.3 which was formulated in Radulović (2004). We feel the need to provide a short explanation that this interesting property holds for any random variables with finite second moments. For the reader's convenience we recall the Proposition 4.3 below.

Proposition 6.1. *For any random variable W, such that $\mathbb{E}W^2 < \infty$, there exists a positive increasing function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ such that*

$$\lim_{x\to\infty} \frac{\phi(x)}{x^2} = +\infty \ \text{ and } \ \mathbb{E}\phi(W) < \infty.$$

Proof. Consider some positive, increasing function $\bar{F}(x)$ such that

$$\epsilon(x) = \lim_{x\to\infty} x^2 \bar{F}(x) = 0.$$

Firstly, we consider the case, when W has bounded support. Let f be a probability density function of W. For some sufficiently large x_0, we put

$$\begin{aligned}\phi(x) &= \begin{cases} \frac{x^2}{\epsilon(x)} & \text{if } \epsilon(x) \neq 0 \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} \frac{1}{1-\bar{F}(x)} & \text{if } \epsilon(x) \neq 0 \\ 1 & \text{else.} \end{cases}\end{aligned}$$

Then, we have

$$\lim_{x\to\infty} \frac{\phi(x)}{x^2} = 0$$

and

$$\begin{aligned}\int_{x_0}^{\infty} \frac{x^2}{\epsilon(x)} f(x)dx &= \int_{x_0}^{\infty} \phi(x)\bar{F}(dx) \\ &= \int_{x_0}^{\infty} \frac{\bar{F}(dx)}{1-\bar{F}(x)} dx < \infty.\end{aligned}$$

For the unbounded support case, just put for some sufficiently large x_0 :

$$\phi(x) = \begin{cases} \frac{x^2}{\epsilon(x)} & \text{if } \epsilon(x) \neq 0 \\ C & \text{else} \end{cases}$$

for some $C > 0$, then the reasoning is going analogously as in the bounded support case. □

References

[1] Barbe, Ph., Bertail, P. (1995). *The Weighted Bootstrap.* Lecture Notes in Statistics, **98**, Springer Verlag, New-York.

[2] Bertail, P., Clémençon, S. (2006). Regenerative block bootstrap for Markov chains. Bernoulli, **12**, 689-712.

[3] Bertail, P., Clémençon , S. (2006). Regeneration-based statistics for Harris recurrent Markov chains. Dependence in Probability and Statistics, Lecture Notes in Statistics, 187, Springer.

[4] Bertail, P., Clémençon, S. (2007). Second-order properties of regeneration-based bootstrap for atomic Markov chains. Test, **16**, 109122

[5] Bertail, P., Clémençon, S. (2015). Bootstrapping roboust statistics for Markovian data applications to regenerative $\mathcal{R}$- and $\mathcal{L}$statistics. Journal of Time Series Analysis, **36**, 462-480.

[6] Bolthausen, E. (1982). The Berry-Essen Theorem for strongly mixing Harris recurrent Markov chains. Z. Wahrsch. Verw. Geb., **60**, 283-289.

[7] Datta, S., McCormick, W. (1995). Some continuous Edgeworth expansions for Markov chains with applications to bootstrap. J. Mult. Analysis, **52**, 83-106.

[8] Dudley, R.M. (1990). Non linear functionals of empirical measures and the bootstrap. Probab. in Banach Spaces, Birkhausen, Boston, **7**, 63-82.

[9] Efron, B. (1979). Bootstrap methods: another look at the jackknife. Ann. Stat., **7**, 1-26.

[10] Giné, E., Zinn, J. (1990). Bootstrapping general empirical measures. Ann. Probab., **18** (1990), 851-869.

[11] Gorst-Rasmussen, A., Bøgsted, M. (2009). Asymptotic Inference for Waiting Times and Patiences in Queues with Abandonment. Communication in Statistics- Simulation and Computation, 01/2009; 38(2), 318-334

[12] Hoffmann-Jørgensen, J. (1991). *Stochastic processes on Polish spaces.* Aarhus universitet, Matematisk institut.

[13] Jain, J., Jamison, B. (1967). Contributions to Doeblin's theory of Markov processes. Z. Wahrsch. Verw. Geb., **8**, 19-40.

[14] Lahiri, S. (2003). *Resampling methods for Dependent Data.* Springer Verlag.

[15] Levental, S. (1988). Uniform limit theorems for Harris recurrent Markov chains. Probab. Theory and Related Fields, 80, 101-118.

[16] Martin R., Yohai, V. (1986). Influence functionals for time series. Ann. Stat., **14**, 781-818.

[17] Meyn, S., Tweedie, R., (1996). *Markov chains and stochastic stability.* Springer.

[18] Nummelin, E. (1978). A splitting technique for Harris recurrent Markov chains. Z. Wahrsch. Verw. Gebiete, **43**, 309-318.

[19] Nummelin, E. (1984). *General Irreducible Markov Chains and Non-negative Operators.* Cambridge Univ. Press, Cambridge.

[20] Radulović, D. (2004). Renewal type bootstrap for Markov chains . Springer, Vol. 13.2004, 1, 147-192.

[21] Rosenblatt, M. (1956). Central Limit Theorem and a strong mixing condition. Ann. Inst. Stat. Math., **42**, 253-268.

[22] Sen B. (2008). *A study of bootstrap and likelihood based methods in nonstandard problems*. A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Statistics). University of Michigan.

[23] Tsai, T.H. (1998). *The uniform CLT and LIL for Markov chains.* Ph.D. thesis, University of Wisconsin.

[24] Rosenthal, J.S. (2006). *First Look at Rigorous Probability Theory.* World Scientific Publishing Co.

[25] van der Vaart A.W., Wellner J.A. (1996). *Weak Convergence and Empirical Processes With Applications to Statistics*. Springer Series in Statistics.

[26] van der Vaart A. W. (2000). *Asymptotic Statistics*. Cambridge University Press.