Electronic Journal of Statistics
Vol. 18 (2024) 4851–4881
ISSN: 1935-7524
https://doi.org/10.1214/24-EJS2318

Regenerative bootstrap for β-null recurrent Markov chains

Carlos A. Fernández[1],

[1]*LTCI, Telecom Paris, Institut Polytechnique de Paris,*
e-mail: fernandez@telecom-paris.fr

Abstract: Two regeneration-based bootstrap methods, namely, the *Regeneration based-bootstrap* [3, 23] and the *Regenerative Block bootstrap* [11] are shown to be valid for the problem of estimating the integral of a function with respect to the invariant measure in a β-null recurrent Markov chain with an accessible atom. An extension of the Central Limit Theorem for randomly indexed sequences is also presented.

MSC2020 subject classifications: Primary 60K35, 60K35; secondary 60K35.
Keywords and phrases: Bootstrap, regeneration based bootstrap, regeneration block bootstrap, nonparametric estimation, null recurrent Markov chain.

Received June 2024.

In [27], Bradley Efron introduced the Bootstrap as a way to overcome some limitations of classical methods that often relied on strong assumptions about the data's underlying distribution or the model's form. Since then, these techniques, first studied in the i.i.d. case, have been developed and extended to time-series (see [46] for an extensive survey of methods) and applied to a wide range of problems in various fields such as signal processing [67, 68], soil science [66] and econometrics [48, 33]. These methods are easy to implement with modern computing power and can provide more accurate and reliable inferences than traditional methods in many situations.

Although originally designed for i.i.d. sampling, there has been significant interest in adapting the bootstrap to situations where the data is dependent. Several resampling methods have been proposed for time series data: these include the autoregressive-sieve bootstrap [43], block bootstrap [45], circular bootstrap [61], the stationary bootstrap [62], continuous-path block bootstrap [56], tapered block bootstrap [57], frequency-domain bootstrap [55, 41], and local bootstrap [59]. For detailed reviews and comparisons of these methods see [28, 42, 40, 17] and the references therein.

In the Markovian case, numerous approaches have been developed and examined. In [44], the authors proposed a block resampling scheme that consists in resampling from a nonparametric estimate of the one-step transition matrix of a finite state Markov chain. This method was extended to the countable case in [3]. Extensions of this method have been proposed for the case where the state

arXiv: 2407.05284

space is Euclidean, as seen in [64], [58, 60] and [32]. The general concept is to estimate the marginal distribution and the transition probability function using a nonparametric function estimation technique and then resample from those estimates. For a detailed explanation of this approach, refer to Section 4 in [40].

A completely new approach to this problem was introduced in [3]. Instead of using estimated transition probabilities, they exploit the regeneration properties of a Markov chain when an accessible atom is visited infinitely often. The main idea underlying this method consists in dividing the chain into a random number of i.i.d. regeneration blocks and then resampling the same number of regeneration blocks. This method, named *Regeneration based bootstrap*, was proved to be valid for finite state atomic chains in [3], and it was extended to general atomic positive recurrent Markov chains in [23].

It was pointed out in [10] that the *Regeneration based bootstrap* is not second-order correct (its rate is $O_{\mathbb{P}}(n^{-1/2})$ only). To overcome this limitation, a variation of this method, called *Regenerative Block bootstrap* (RBB), was introduced in [11]. This method consists in imitating the renewal structure of the chain by sampling regeneration data blocks, until the length of the reconstructed bootstrap series is larger than the length n of the original data series (notice the difference with the *Regeneration based bootstrap*, where the number of sampled blocks is equal to the number of regeneration blocks in the original chain). It was shown in [11] that, for atomic positive recurrent Markov chains, the RBB for estimating the integral of a function with respect to the invariant probability, has a uniform rate of convergence of order $O_{\mathbb{P}}\left(n^{-1}\right)$ (the same as in the i.i.d. case).

Despite all these efforts in the positive recurrent case, up to our knowledge, no bootstrap method has been studied in the general null-recurrent scenario, although some specific AR(1) models with unit roots have been the subject of investigations. It has been shown that, for those AR(1) models the standard bootstrap methods (including parametric ones) do not work [22, 54] unless one works under the null hypothesis of unit root (see [8]). This idea can not be applied to general null recurrent Markov chain unless one specifies the parametric or semiparametric forms of the Markov chain. Hence, our objective in this paper is to propose a general valid method and show that, provided that the estimators are correctly standardized, both *Regeneration based-bootstrap* and *Regenerative Block bootstrap* are valid schemes for estimating integrals with respect to the invariant measure when the Markov chain is β-null recurrent and possesses an accessible atom. The task is challenging because the expectation of the time of return to an atom of such Markov chains is infinite, the bootstrap of such quantity does not work [4, 39]: indeed a necessary condition for the bootstrap to work is generally the finiteness of the variance [21]. Despite this fact, we will show that, by using the correct random normalization, one can obtain a CLT for the mean and the validity of both bootstrap regeneration methods in a null-recurrent framework.

The main difference between the methods described here and its counterparts in the positive recurrent case is the fact that we use renormalization based on the number of regenerations blocks (denoted by $T(n)$) instead of n. Given that

the number of blocks is of order much smaller than n, it makes the method more restrictive. But to our knowledge, there are no other bootstrap methods available in this null-recurrent context (unless we consider a specific parametric model). Notice that, even standard estimators will have rate of convergence $\sqrt{T(n)}$ with $T(n)$ or order n^β (see [37]) which of course also limits its applicability. For symmetric random walks, $\beta = 1/2$, it is known that estimators of linear functionals, kernels estimators, volatility estimators have rate close to $n^{-1/4}$, [24].

In order to make the exposition simpler, our asymptotic results will be specifically stated for integrals with respect to the invariant measure, however, the procedures can be applied to any statistic defined over a regeneration blocks, as long as it has finite variance. Similarly, our results are stated for first order Markov chains, but they can easily be extended to higher order chains by vectorization [26, pp. 15].

The paper is organized as follows: in section 1 we provide a brief introduction to null recurrent Markov chains, making a special emphasis on atomic ones and presenting the main results that we use throughout the paper. In subsection 1.3 we present an extension of the Central Limit Theorem for randomly indexed sequences (Lemma 1.2). Section 2 is dedicated to the *Regenerative Block bootstrap* in β-null recurrent Markov chains, while Section 3 is devoted to the *Regeneration based-bootstrap*. In section 4 we have performed two simulation studies to show the behavior of both algorithms in practice. Section 5 contains a few concluding remarks. The technical proofs are postponed to Section 6.

1. A short introduction to null-recurrent Markov chains

In this section, we introduce some notation and review some important concepts from Markov chain theory that will be used throughout the paper. For more details, please refer to [50, 26].

1.1. Notation and definitions

Consider an homogeneous Markov chain $\mathbf{X} = \{X_0, X_1, \ldots\}$ on a countably generated state space $(E, \mathcal{E})$, with transition kernel P and initial probability distribution λ. This means that for any $B \in \mathcal{E}$ and $n \in \mathbb{N}$, we have $\mathcal{L}(X_0) = \lambda$ and

$$\mathbb{P}(X_{n+1} \in B \mid X_0, \ldots, X_n) = P(X_n, B) \quad \text{almost surely.}$$

Note that the assumption of a countably generated state space is commonly used in Markov chain theory to avoid pathological examples known as 'anormal' chains [25]. For more information on this topic, see [25], [35], and [65]. An example of an 'anormal' chain can be found in [15]. This assumption does not significantly limit the generality of our results since most of the time $\mathcal{E}$ will be the borelian σ-algebra of $\mathbb{R}^d$, which is countably generated.

In the following, we use $\mathbb{P}_\lambda$ (or $\mathbb{P}_x$ for x in E) to denote the probability measure on the underlying space such that $X_0 \sim \lambda$ (or $X_0 = x$). We use $\mathbb{E}_\lambda$ to

represent the $\mathbb{P}_\lambda$-expectation (or $\mathbb{E}_x$ to represent the $\mathbb{P}_x$-expectation), and $\mathbb{I}\{\mathcal{A}\}$ to represent the indicator function of event $\mathcal{A}$.

A homogeneous Markov chain is said to be irreducible if there exists a σ-finite measure ϕ on $(E, \mathcal{E})$ such that for all $x \in E$ and all $A \in \mathcal{E}$ with $\phi(A) > 0$, there exists some $n \geqslant 1$ such that $P^n(x, A) > 0$. In this case, there exists a maximal irreducibility measure ψ with respect to which all other irreducibility measures are absolutely continuous. If $\mathbf{X}$ is ψ-irreducible, there is $d' \in \mathbb{N}^*$ and disjoints sets $D_1, \ldots, D_{d'}$ $D_{d'+1} = D_1$ weighted by ψ such that $\psi(E \backslash \cup_{1 \leqslant i \leqslant d'} D_i) = 0$ and $\forall x \in D_i, P(x, D_{i+1}) = 1$. The g.c.d. d of such integers is called the *period* of the chain. $\mathbf{X}$ is said to be *aperiodic* if $d = 1$.

Thorough this paper, we assume that the Markov chains under consideration are homogeneous, aperiodic, and irreducible with maximal irreducibility measure ψ.

An irreducible chain possesses an accessible atom, if there is a set $\boldsymbol{\alpha} \in \mathcal{E}$ such that for all x, y in $\boldsymbol{\alpha}$: $P(x, \bullet) = P(y, \bullet)$ and $\psi(\boldsymbol{\alpha}) > 0$. For instance, when a chain can take a countable number of values, any single point visited by the chain is an atom. Denote by $\sigma_{\boldsymbol{\alpha}}$ and $\tau_{\boldsymbol{\alpha}}$, respectively, the times of first visit and first return of the chain to $\boldsymbol{\alpha}$, i.e. $\tau_{\boldsymbol{\alpha}} = \inf\{n \geqslant 1 : X_n \in \boldsymbol{\alpha}\}$ and $\sigma_{\boldsymbol{\alpha}} = \inf\{n \geqslant 0 : X_n \in \boldsymbol{\alpha}\}$. The subsequent visit and return times $\sigma_{\boldsymbol{\alpha}}, \tau_{\boldsymbol{\alpha}}(k)$, $k \geqslant 1$ are defined inductively as follows:

$$\tau_{\boldsymbol{\alpha}}(1) = \tau_{\boldsymbol{\alpha}} \quad , \quad \tau_{\boldsymbol{\alpha}}(k) = \min\{n > \tau_{\boldsymbol{\alpha}}(k-1) : X_n \in \boldsymbol{\alpha}\}, \tag{1}$$

$$\sigma_{\boldsymbol{\alpha}}(1) = \sigma_{\boldsymbol{\alpha}} \quad , \quad \sigma_{\boldsymbol{\alpha}}(k) = \min\{n > \sigma_{\boldsymbol{\alpha}}(k-1) : X_n \in \boldsymbol{\alpha}\}. \tag{2}$$

We use $T_n(A)$ to represent the random variable that counts the number of times the chain visits the set A up to time n, i.e., $T_n(A) = \sum_{t=0}^{n} \mathbb{I}\{X_t \in A\}$. Similarly, we use $T_\infty(A)$ to represent the total number of visits of chain $\mathbf{X}$ to A.

An atom $\boldsymbol{\alpha}$ is called *recurrent* if $\mathbb{E}_x T_\infty(\boldsymbol{\alpha}) = +\infty$ for all $x \in \boldsymbol{\alpha}$; otherwise, it is called *transient*. A notable property of atomic chains is that all accessible atoms are either all recurrent or all transient. Therefore, we say that an atomic chain is recurrent if one (and thus all) of its accessible atoms is recurrent. If $\mathbf{X}$ is aperiodic, recurrent and possesses an accessible atom, then the probability of returning infinitely often to the atom $\boldsymbol{\alpha}$ is equal to one, no matter the starting point, i.e.

$$\mathbb{P}_x\big(T_\infty(\boldsymbol{\alpha}) = \infty\big) = 1 \quad \forall x \in E.$$

Denote by $\mathbb{P}_{\boldsymbol{\alpha}}$ and $\mathbb{E}_{\boldsymbol{\alpha}}$ the probability and the expectation conditionally to $X_0 \in \boldsymbol{\alpha}$.

A fundamental tool for understanding the long-term behavior of Markov chains is the existence of invariant measures, that is, a measure π such that

$$\pi(A) = \int P(x, A)\, d\pi(x) \quad \forall A \in \mathcal{E}.$$

Every irreducible and recurrent Markov chain admits a unique (up to a multiplicative constant) invariant measure [50, Theorem 10.4.9]. In the atomic

case, the invariant measure is just the occupation measure over the first block $\mathcal{B}_1 = \left(X_{\tau_{\boldsymbol{\alpha}}(1)+1}, \ldots, X_{\tau_{\boldsymbol{\alpha}}(2)}\right)$ [26, Theorem 6.4.2], i.e.

$$\pi_{\boldsymbol{\alpha}}(A) = \mathbb{E}_{\boldsymbol{\alpha}}\left(\sum_{j=1}^{\tau_{\boldsymbol{\alpha}}} \mathbb{I}\{X_j \in A\}\right), \quad \forall A \in \mathcal{E}. \tag{3}$$

An irreducible Markov chain is *positive recurrent* if its invariant measure is finite. When the invariant measure is just σ-finite, then the chain is called *null recurrent*. From (3), it is clear that an atomic Markov chain is positive recurrent if and only if $\mathbb{E}_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}} < +\infty$, and in this case, the measure defined by $\pi_{\boldsymbol{\alpha}}/\mathbb{E}_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}}$ is an invariant probability for the chain. The existence of this invariant probability makes the theory of positive recurrent Markov chains, very similar to the i.i.d. case [50, Chapter 17].

Conversely, dealing with null recurrent chains is considerably more challenging, and a comprehensive theory of non-parametric estimation for this type of chain does not exist. To address this issue, Karlsen and Tjøstheim introduced in [37] a regularity condition for the tail behavior of the distribution of $\tau_{\boldsymbol{\alpha}}$ that renders the problem more tractable. Specifically, denote by Γ the gamma function[1], then, a chain is referred to as β-null recurrent if there is a constant $\beta \in (0,1)$ and a slowly varying function[2] L such that

$$\mathbb{P}_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}} > n) \sim \frac{1}{\Gamma(1-\beta)n^{\beta}L(n)}. \tag{4}$$

The number β, also known as the *regularity index* (see [18, 19]) satisfies

$$\beta = \sup\{p > 0 : \mathbb{E}_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}}^{p}) < +\infty\}.$$

Some of the most well-known examples of β-null recurrent Markov chains are the random walks in $\mathbb{R}$, which are 1/2-null recurrent [36], the Bessel random walks [20, 1] and some types of threshold autoregressive (TAR) [29] and vector autoregressive processes (VAR) [51]. β-null recurrent Markov chains appear naturally in many fields of statistics and probability for instance for studying population dynamics, statistical mechanics or the study of Polymer.

1.2. Renewal properties and Block decomposition

The *strong Markov property* implies that the sample paths of an atomic Markov chain can be partitioned into independent blocks of random length corresponding to consecutive visits to $\boldsymbol{\alpha}$, given by:

$$\mathcal{B}_0 = \left(X_0, X_1, \ldots, X_{\tau_{\boldsymbol{\alpha}}(1)}\right)$$

[1]The Γ function is defined as $\Gamma(x) = \int_0^{+\infty} t^{x-1}\exp(-t)\,dt$.

[2]A measurable and positive function L is said to be *slowly varying at* $+\infty$ if it is defined in $[a, +\infty)$ for some $a \geqslant 0$, and satisfies $\lim_{x\to+\infty} L(xt)/L(x) = 1$ for all $t \geqslant a$. For a detailed discussion on these types of functions, refer to [14].

$$\begin{aligned}\mathcal{B}_1 &= \left(X_{\tau_{\boldsymbol{\alpha}}(1)+1}, \ldots, X_{\tau_{\boldsymbol{\alpha}}(2)}\right)\\ &\ldots\\ \mathcal{B}_n &= \left(X_{\tau_{\boldsymbol{\alpha}}(n)+1}, \ldots, X_{\tau_{\boldsymbol{\alpha}}(n+1)}\right)\\ &\ldots\end{aligned}$$

blocks $\{\mathcal{B}_j\}_{j\geqslant 1}$ is called *regeneration blocks*. As customary in the β-null recurrent Markov chain literature, we will use $T(n)$ to denote the number of complete regeneration blocks up to time n, i.e. $T(n) = \max\left(T_n(\boldsymbol{\alpha}) - 1, 0\right)$. We will denote by $\ell(\mathcal{B}_i)$ the length of the i-th block, therefore,

$$\ell(\mathcal{B}_j) = \begin{cases} \tau_{\boldsymbol{\alpha}} & , \quad j = 0 \\ \tau_{\boldsymbol{\alpha}}(j+1) - \tau_{\boldsymbol{\alpha}}(j) & , \quad j \geqslant 1 \end{cases} \tag{5}$$

The random variable $T(n)$, and its relationship with $\sum_{j=0}^{k} \ell(\mathcal{B}_j)$, is crucial in the theory we will develop in this paper, therefore, we will state in this section its main properties in the β-null recurrent scenario.

Assume $\mathbf{X}$ is a β-null recurrent Markov chain with an accessible atom $\boldsymbol{\alpha}$. By (3.27) in [37], the function L in (4) can be normalized in such a way that

$$u(z) = z^{\beta} L(z) \tag{6}$$

is a continuous function that is strictly increasing in the interval $[z_0, +\infty)$ for some $z_0 \in \mathbb{R}_+$. Define $v(z)$ as

$$v(z) = u^{(-1)}(z) = \inf\left\{s : u(s) > z\right\}, \tag{7}$$

then, $u(v(z)) = v(u(z)) = z$ for $z \geqslant z_0$.

Consider the space of càdlàg functions defined on the interval $[0, +\infty)$, denoted by $\mathscr{D}_{[0,+\infty)}$. This space consists of the real functions that are right-continuous with left limits and defined over $[0, +\infty)$. More precisely, a function $g \in \mathscr{D}_{[0,+\infty)}$ if and only if g is right-continuous, has left limits at all points $t > 0$, and $\lim_{t\downarrow 0} g(t) = g(0)$. The space $\mathscr{D}_{[0,+\infty)}$ is equipped with the Skorokhod[3] topology, making it a complete and separable metric space. We will use $\xrightarrow{\mathscr{D}_{[0,+\infty)}}$ to denote weak convergence in this space, and $\xrightarrow{\text{fd}}$ for convergence of finite-dimensional laws. Two stochastic processes Y_n, Z_n in $\mathscr{D}_{[0,+\infty)}$ are said to be *equivalent* if $Y_n - Z_n$ converges weakly to the zero process. If $Y_n \xrightarrow{\mathscr{D}_{[0,+\infty)}} Y$ and Y_n and Z_n are equivalent, then $Z_n \xrightarrow{\mathscr{D}_{[0,+\infty)}} Y$ (see Lemma 3.31 in [34]).

Define the following processes

$$T_n(t) = \frac{T(\lfloor nt \rfloor)}{u(n)}, \quad C_n(t) = \frac{1}{v(n)} \sum_{k=0}^{\lfloor nt \rfloor} \ell(\mathcal{B}_k), \tag{8}$$

and $C_n^{(-1)}(t) = \inf\{x : C_n(x) > t\}$. The following Theorem, proved in [37], shows that these three processes converge in $\mathscr{D}_{[0,+\infty)}$ and that T_n and $C_n^{(-1)}$ are equivalent.

[3] See Chapter 6 of [34] or Chapter 3 in [13] for more details about this space.

Theorem 1.1. *Assume $\boldsymbol{X}$ is a β-null recurrent atomic Markov chain. Then,*

i) $C_n \xrightarrow{\mathscr{D}_{[0,+\infty)}} S_\beta$ *where S_β is the one-sided stable Levy process defined by the marginal characteristics*

$$\mathbb{E}\left[\exp\left(isS_\beta(t)\right)\right] = \exp\left(is^\beta t\right) \quad s \in \mathbb{R}, t \in [0,+\infty].$$

ii) $C_n^{(-1)}$ *and T_n are equivalent processes and both converge in $\mathscr{D}_{[0,+\infty)}$ to the Mittag-Leffler process of parameter β.*

Remark 1.1. The Mittag-Leffler process with parameter β is defined as the inverse of S_β. It is a strictly increasing continuous stochastic process defined as

$$M_\beta(t) = t^\beta M_\beta(1) \quad , \quad \mathbb{E}\left(M_\beta^m(1)\right) = \frac{m!}{\Gamma(1+m\beta)} \quad m \geqslant 0.$$

Theorem 1.1 shows a striking difference between positive and null recurrent Markov chains. While in the former the existence of moments for $\ell(B_j)$ implies that C_n and T_n (taking $u(n) = n$) converge almost surely respectively to $t\mathbb{E}_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}}$ and $t/\mathbb{E}_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}}$, and therefore, $T(n)$ can be approximated almost surely by the deterministic quantity $n/\mathbb{E}_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}}$, in the latter, we only have weak convergence, hence $T(n)$ can only be controlled by the deterministic quantity $u(n)$ in distribution.

1.3. Properties of linear functionals defined on β-null recurrent chains

For a measurable function $f : E \to \mathbb{R}$, and an atomic Markov chain $\mathbf{X}$ with an accessible atom $\boldsymbol{\alpha}$, consider the problem of estimating $\pi_{\boldsymbol{\alpha}}(f) = \int f d\pi_{\boldsymbol{\alpha}}$, where $\pi_{\boldsymbol{\alpha}}$ is as in (3) and $\pi_{\boldsymbol{\alpha}}(f) < +\infty$. Denote by $S_n(f)$ the partial sums of f over the chain, that is

$$S_n(f) = \sum_{k=0}^{n} f(X_k). \tag{9}$$

The Ratio Limit Theorem for atomic chains [26, Theorem 6.6.2] shows that if g is a measurable function, then, for every invariant measure π we have

$$\frac{S_n(f)}{S_n(g)} \xrightarrow{\text{a.s.}} \frac{\pi(f)}{\pi(g)}, \tag{10}$$

as long as $\pi(g) \neq 0$.

Remark 1.2. From (10) is clear that $S_n(f)/T(n)$ is a strongly consistent estimator of $\pi_{\boldsymbol{\alpha}}(f)$, and, in the positive recurrent case, $S_n(f)/n \xrightarrow{\text{a.s.}} \pi_{\boldsymbol{\alpha}}(f)/\mathbb{E}_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}}$. In the null recurrent case, however, $S_n(f)/n \xrightarrow{\text{a.s.}} 0$ (see Corollary 6.6.3 in [26]) and there is no deterministic sequence $a(n)$ such that $S_n(f)/a(n)$ converges almost surely to a non-zero limit [18].

Given that our interest in this paper is to apply the bootstrap method to the study of $\pi_{\boldsymbol{\alpha}}(f)$ we need to find a series of i.i.d. random variables whose mean strongly converges to $\pi_{\boldsymbol{\alpha}}(f)$. To do this, define the following random variables

$$f(\mathcal{B}_j) = \begin{cases} \sum_{i=0}^{\tau_\alpha} f(X_i) & , \quad j = 0 \\ \sum_{i=\tau_{\boldsymbol{\alpha}}(j)+1}^{\tau_\alpha(j+1)} f(X_i) & , \quad j \geqslant 1 \end{cases}.$$

The strong Markov property implies that under $\mathbb{P}_{\boldsymbol{\alpha}}$, the sequence $\{f(\mathcal{B}_j)\}_{j\geqslant 0}$ is i.i.d. Moreover, for every initial probability λ such that $\mathbb{P}_\lambda(\tau_{\boldsymbol{\alpha}} < \infty) = 1$, the random variables $f(\mathcal{B}_j), j \geqslant 0$ are independent and for $j \geqslant 1$ they are i.i.d. Therefore, $S_n(f)$ can now be written as a sum of independent random variables as follows:

$$S_n(f) = f(\mathcal{B}_0) + \sum_{j=1}^{T(n)} f(\mathcal{B}_j) + \sum_{i=\tau_{\boldsymbol{\alpha}}(T(n)+1)+1}^{n} f(X_i), \tag{11}$$

with the convention that the sum of an empty set is 0. As customary in the β-null recurrent literature, we will denote the last term in (11) by $f(\mathcal{B}_{(n)})$.

Equation (3) indicates that $\mathbb{E}_{\boldsymbol{\alpha}} f(\mathcal{B}_j) = \pi_{\boldsymbol{\alpha}}(f)$ for $j \geqslant 1$, hence, if we assume that $\pi_{\boldsymbol{\alpha}}(|f|) < +\infty$, the Law of Large Numbers for randomly indexed sequences [30, Theorem 8.2, pp 302] shows that

$$\frac{1}{T(n)} \sum_{j=1}^{T(n)} f(\mathcal{B}_j) \xrightarrow{\text{a.s.}} \pi_{\boldsymbol{\alpha}}(f). \tag{12}$$

Remark 1.3. The recurrence of the chain implies that $T(n) \to \infty$ almost surely, therefore $f(\mathcal{B}_0)/T(n)$ and $f(\mathcal{B}_{(n)})/T(n)$ converge to 0 almost surely (see Lemma 1 in [6]). This allows us to consider only the i.i.d. blocks $f(\mathcal{B}_j), j \geqslant 1$ in our estimations.

If we suppose further that $f(\mathcal{B}_1)$ has finite second moment, and we denote by σ^2 the variance of $f(\mathcal{B}_1)$, then

$$\widehat{\sigma}_n^2 = \frac{1}{T(n)} \sum_{j=1}^{T(n)} \left(f(\mathcal{B}_j) - \frac{1}{T(n)} \sum_{i=1}^{T(n)} f(\mathcal{B}_i) \right)^2 \xrightarrow{\text{a.s.}} \sigma^2. \tag{13}$$

Much of the work carried out in this investigation deals with sequences indexed by the sequence of random variables $T(n)$. As explained at the end of Section 1.2, this sequence, although it converges almost surely to $+\infty$, can not be deterministically approximated in probability, it only admits an approximation in distribution. This creates huge problems, even for simple tasks, as to obtaining a CLT, because CLTs for randomly indexed sequences (see [2] for the original formulation and Th. 17.2 in [13] for its more general form) require

being able to control deterministically, at least in probability, the sequence of the number of terms. The result we present below extends this CLT, replacing the requirement of the control in probability by the existence of the limit of a stochastic process defined in terms of the sequence of the number of terms.

Lemma 1.2 (CLT for randomly indexed sequences)**.** *Let $X_1, X_2 \dots$ be i.i.d. random variables such that $\mathbb{E}(X_1) = \mu$ and $\operatorname{Var} X_1 = \sigma^2 > 0$. Let $N(n)$ be a sequence of integer-valued random variables. Assume there exists an unbounded increasing sequence of real numbers u_n such that the process $N_n(t) = N(\lfloor nt \rfloor)/u_n$ satisfies the following conditions:*

- *There exists a process S_n in $\mathscr{D}_{[0,+\infty)}$ such that, for each n it is non-negative and non-decreasing.*
- *$S_n \xrightarrow{\mathscr{D}_{[0,+\infty)}} S$ where S is a strictly increasing non-negative process with independent increments, no fixed jumps, and $S(0) \equiv 0$.*
- *N_n is equivalent to $S_n^{(-1)}$.*

Then, N_n converges to $S^{(-1)}$ in $\mathscr{D}_{[0,+\infty)}$,

$$\frac{\sqrt{N(n)}}{\sigma}\left(\frac{1}{N(n)}\sum_{j=1}^{N(n)}(X_j-\mu)\right) \xrightarrow{d} \boldsymbol{N}(0,1),$$

and $N_n(1)$ and $\frac{\sqrt{N(n)}}{\sigma}\left(\frac{1}{N(n)}\sum_{j=1}^{N(n)}(X_j-\mu)\right)$ are asymptotically independent.

Corollary 1.1. [Theorem 17.2 in [13]] Suppose $X_1, \dots, X_n$ are i.i.d. with $\mathbb{E}X_1 = \mu$ and $\operatorname{Var} X_1 = \sigma^2$. If $N(n)$ is a sequence of integer-valued random variables such that

$$\frac{N(n)}{u_n} \xrightarrow{p} \theta, \tag{14}$$

where θ is a positive random variable and u_n is a sequence of positive numbers going to infinity, then

$$\frac{\sqrt{N(n)}}{\sigma}\left(\frac{1}{N(n)}\sum_{j=1}^{N(n)}(X_j-\mu)\right) \xrightarrow{d} \mathbf{N}(0,1).$$

Using Lemma 1.2 and Theorem 1.1 we can provide a different proof of the following Central Limit Theorem for β-null recurrent atomic Markov chains, which was originally proved in [5].

Theorem 1.3. *Let $\boldsymbol{X}$ be a β-null recurrent Markov chain, with an accessible atom $\boldsymbol{\alpha}$. For every $\pi_{\boldsymbol{\alpha}}$- measurable function f such that $\sigma^2 = \operatorname{Var} f(\mathcal{B}_1)$ is finite, we have the following convergence in distribution:*

$$\frac{\sqrt{T(n)}}{\sigma}\left(\frac{1}{T(n)}\sum_{j=1}^{T(n)} f(\mathcal{B}_j) - \int f\, d\pi_{\boldsymbol{\alpha}}\right) \xrightarrow{d} \boldsymbol{N}(0,1). \tag{15}$$

Moreover, $T(n)/n^{\beta}L(n)$ converges to a Mittag-Leffler distribution with parameter β and it is asymptotically independent of the left-hand side of (15). If in addition we also have $\mathbb{E}[|f|(\mathcal{B}_1)^2] < +\infty$, then

$$\frac{\sqrt{T(n)}}{\sigma}\left(\frac{S_n(f)}{T(n)} - \int f\, d\pi_{\boldsymbol{\alpha}}\right) \xrightarrow{d} \boldsymbol{N}(0,1). \tag{16}$$

The following corollary is a direct consequence of Theorem 1.3, equation (13) and Slutsky's theorem.

Corollary 1.2. Under the same hypothesis of Theorem 1.3, we have

$$\frac{\sqrt{T(n)}}{\widehat{\sigma}_n}\left(\frac{1}{T(n)}\sum_{j=1}^{T(n)} f(\mathcal{B}_j) - \int f\, d\pi_{\boldsymbol{\alpha}}\right) \xrightarrow{d} \mathbf{N}(0,1),$$

and if $\mathbb{E}[|f|(\mathcal{B}_1)^2] < +\infty$ also holds, then

$$\frac{\sqrt{T(n)}}{\widehat{\sigma}_n}\left(\frac{S_n(f)}{T(n)} - \int f\, d\pi_{\boldsymbol{\alpha}}\right) \xrightarrow{d} \mathbf{N}(0,1).$$

2. The regenerative block-bootstrap algorithm (RBB)

Let $\mathbf{X}^{(n)} = (X_0, ..., X_n)$ be observations drawn from a β-null recurrent Markov chain $\mathbf{X}$ with an *a priori* known accessible atom $\boldsymbol{\alpha}$. As in the previous section, let f be a $\pi_{\boldsymbol{\alpha}}$-integrable function such that $f(\mathcal{B}_1)$ has a finite second moment. Denote by σ^2 the variance of $f(\mathcal{B}_1)$.

The Regenerative block-bootstrap (RBB) method, which we explore in this section, was initially introduced in [11] for positive recurrent Markov chains. In their Theorem 2.1, it was shown that, in the atomic case, the RBB distribution achieves a uniform rate of convergence of order $O_p\left(n^{-1}\right)$ for both the studentized and unstudentized sample mean, meaning that the sup-norm between the true distribution and its bootstrap approximation is of order $O_p\left(n^{-1}\right)$.

In this section, we show that the method is also applicable in the β-null recurrent case, although we have not been able to obtain a rate.

Remark 2.1. Obtaining rates of convergence, for the bootstrap, typically depends on Edgeworth expansions [31]. These expansions can be derived, at least formally, by calculating cumulants using standard techniques. In the Markovian case, the validity of these expansions not only depends on the cumulants of $f(\mathcal{B}_j)$ but also on the moments of $\tau_{\boldsymbol{\alpha}}$ [49, 9]. More precisely, up to our knowledge, these expansions have been obtained only when $\mathbb{E}_{\boldsymbol{\alpha}}\tau_{\boldsymbol{\alpha}}^4$ is finite [9, Theorem 5.1]. Developing methods to obtain Edgeworth expansions for distributions with very few moments is an interesting research direction, but it would involve substantial theoretical developments that are beyond the scope of this work.

Proposition 3.1 in [9] shows that for positive recurrent chains, in the nonstationary case (when the initial law λ is not the invariant probability measure),

the first data block $\mathcal{B}_0$ induces a bias of order $O(n^{-1})$, which cannot be estimated from a single realization $\mathbf{X}^{(n)}$ of the chain starting from λ. The last block $\mathcal{B}_{(n)}$ (which is incomplete) induces a first-order term in the bias too. This led the authors in [11] to only consider statistics based on the regenerative data blocks $\mathcal{B}_1, \ldots, \mathcal{B}_{T(n)}$.

In the β-null recurrent case, the lack of finite first moment for the block sizes suggests that considering the non-regenerative blocks will incur in an even worse bias, hence, as in [11], we will only consider statistics based on the regenerative data blocks $\mathcal{B}_1, \ldots, \mathcal{B}_{T(n)}$.

While our asymptotic results are specifically stated for integrals with respect to the invariant measure, the algorithm can be applied to any statistic defined over the regeneration blocks, as long as it has finite variance.

As customary in the bootstrap literature, $\mathbb{P}^*(\bullet) = \mathbb{P}(\bullet \mid \mathbf{X}^{(n)})$ denotes the conditional probability given $\mathbf{X}^{(n)}$. We will write $Z_n^* \xrightarrow{d^*}_p Z$ to indicate the weak converge in probability of the bootstrap random variables Z_n^* to Z, this is, for all $x \in \mathbb{R}$, $\mathbb{P}^*(Z_n^* \leqslant x) \xrightarrow{p} \mathbb{P}(Z \leqslant x)$. See pp. 2550 in [16] for more details.

In this section, our goal is to bootstrap a general statistic G_n that converges to a parameter θ. We will typically prove asymptotic results for the case where $G_n = \sum_{j=1}^{T(n)} f(\mathcal{B}_i)/T(n)$. Additionally, we assume the availability of a block-based standardization, denoted as $Std_n = Std(\mathcal{B}_1, ..., \mathcal{B}_{T(n)})$. The distribution of interest is defined as $H_n(x) = \mathbb{P}(Std_n^{-1}(G_n - \theta) \leqslant x)$.

The RBB procedure is performed in four steps as follows:

1. Count the number of visits $T_n(\boldsymbol{\alpha})$ to the atom $\boldsymbol{\alpha}$ up to time n, and divide the observed sample path $\mathbf{X}^{(n)} = (X_0, \ldots, X_n)$ into $T_n(\boldsymbol{\alpha}) + 1$ blocks, $\mathcal{B}_0$, $\mathcal{B}_1, \ldots, \mathcal{B}_{T_n(\boldsymbol{\alpha})-1}$, $\mathcal{B}^{(n)}_{T_n(\boldsymbol{\alpha})}$, corresponding to the pieces of the sample path between consecutive visits to the atom $\boldsymbol{\alpha}$. Drop the first and last (non-regenerative) blocks. Denote by $T(n)$ the number of remaining blocks.
2. Draw sequentially bootstrap data blocks $\mathcal{B}^*_{1,T(n)}, \ldots, \mathcal{B}^*_{k,T(n)}$ independently from the empirical distribution $F_n = T(n)^{-1} \sum_{j=1}^{T(n)} \delta_{\mathcal{B}_j}$ of the blocks $\{\mathcal{B}_j\}_{1 \leqslant j \leqslant T(n)}$ conditioned on $\mathbf{X}^{(n)}$, until the length of the bootstrap data series, $\ell^*(k) = \sum_{j=1}^{k} \ell(\mathcal{B}^*_{j,T(n)})$, is larger than n. Let $T_n^*(\boldsymbol{\alpha}) = \inf\{k \geqslant 1, \ell^*(k) > n\}$ and $T^*(n, T(n)) = T_n^*(\boldsymbol{\alpha}) - 1$.
3. From the data blocks generated in step 2, reconstruct a pseudo-trajectory of size $\ell^*(T^*(n, T(n)))$ by binding the blocks together, that is

$$X^{*(n)} = \left(\mathcal{B}^*_{1,T(n)}, ..., \mathcal{B}^*_{T^*(n,T(n)),T(n)}\right).$$

   Compute the *RBB statistic* $G_n^* = G_n(X^{*(n)})$.
4. If $Std_n = S(\mathcal{B}_1, ..., \mathcal{B}_{T(n)})$ is an appropriate standardization of the original statistic G_n, compute $Std_n^* = S(\mathcal{B}^*_{1,T(n)}, ..., \mathcal{B}^*_{T^*(n,T(n)),T(n)})$.

The *RBB distribution* is then given by

$$H_{RBB}(x) = \mathbb{P}^*\left(Std_n^{*-1}(G_n^* - G_n) \leqslant x\right).$$

One purpose of the next paragraphs is to show that if we choose a correct standardization Std_n, then we can obtain that $H_{RBB}(x) - H_n(x) \xrightarrow{p} 0$ uniformly in x. Our main asymptotic result, in the case of integrals concerning the invariant measure, is the following.

Theorem 2.1 (Validity of the RBB)**.** *Let $\boldsymbol{X}$ be a β-null recurrent Markov chain with an accessible atom $\boldsymbol{\alpha}$, and let f be a $\pi_{\boldsymbol{\alpha}}$-integrable function such that $\mathbb{E}[f(\mathcal{B}_1)^2] < +\infty$. Define*

$$\widehat{\sigma}^2_{T(n)} = \frac{1}{T(n)} \sum_{j=1}^{T(n)} \left(f(\mathcal{B}_j) - \frac{1}{T(n)} \sum_{i=1}^{T(n)} f(\mathcal{B}_i) \right)^2 \text{ and } \widehat{\mu}_{T(n)} = \frac{1}{T(n)} \sum_{i=1}^{T(n)} f(\mathcal{B}_i).$$

Then we have,

$$\frac{\sqrt{T^*(n, T(n))}}{\widehat{\sigma}_{T(n)}} \left(\frac{1}{T^*(n, T(n))} \sum_{j=1}^{T^*(n,T(n))} \left(f(\mathcal{B}^*_{j,T(n)}) - \widehat{\mu}_{T(n)} \right) \right) \xrightarrow{d^*}_p \boldsymbol{N}(0,1)\,.$$

This theorem yields that the bootstrap distribution of the standardized sum has asymptotically the same distribution as the statistics $\sum_{j=1}^{T(n)} f(\mathcal{B}_j)/T(n)$ estimating $\int f d\pi_{\boldsymbol{\alpha}}$. The proof of this result is non-trivial and totally non-standard: it starts by constructing a space, via Skorokhod-Dudley-Wichura Theorem (see pp. 1171 in [39]), in which we can get a.s. convergence of order statistics of the block lengths, as in [39]. Then, in that space we apply the CLT described in Lemma 1.2 to obtain the convergence in probability of the bootstrap quantity H_{RBB} to the CDF of a normal distribution, which implies convergence of the same things in distribution in the original space. But since this bootstrap limit is non-random (it does not depend on the data), we get in turn the weakly convergence in probability. The regenerative block bootstrap is thus first-order correct. In particular, this justifies the use of the quantiles of the bootstrap distribution (with or without standardizing) to obtain confidence intervals for $\int f d\pi_{\boldsymbol{\alpha}}$.

Remark 2.2. In the original formulation of the RBB for atomic and positive recurrent chains [11, Theorem 2.1], the estimator used was $G_n = \sum_{i=1}^{T(n)} f(\mathcal{B}_i)/n_{\boldsymbol{\alpha}}$, where $n_{\boldsymbol{\alpha}} = \sum_{k=1}^{T(n)} \ell(\mathcal{B}_k)$. A key element in their proof is that $n_{\boldsymbol{\alpha}}$ is $a.s.$ equivalent to a multiple of n, however, in the null-recurrent scenario, this equivalence does not hold due to the lack of first moment for $\ell(\mathcal{B}_1)$. Therefore, we need to use the random normalization. On the other hand, Remarks 1.2 and 1.3 rule out the use of $\sum_{i=1}^{T(n)} f(\mathcal{B}_i)/n$ in the null-recurrence case (it converges $a.s.$ to 0), and equation (12) suggests $\sum_{j=1}^{T(n)} f(\mathcal{B}_i)/T(n)$ as its natural replacement. It should be pointed out that using $\sum_{i=1}^{T(n)} f(\mathcal{B}_i)/u(n)$ (or $S_n(f)/u(n)$) is also not useful, because its limit distribution is a constant multiple of a Mittag-Leffler distribution, see [18, Theorem 2.1].The random normalization seems unavoidable in the β-null recurrent scenario.

3. The regeneration-based bootstrap algorithm

In this section, we adapt the *Regeneration-based bootstrap* to the β-null recurrent Markov chain scenario.

Similarly to Section 2, consider observations $\mathbf{X}^{(n)} = (X_0, \ldots, X_n)$ drawn from a β-null recurrent Markov chain $\mathbf{X}$ that has an accessible atom $\boldsymbol{\alpha}$ known beforehand. Suppose that f is a function such $\pi_{\boldsymbol{\alpha}}(f)$ is finite and the second moment of $f(\mathcal{B}_1)$ is also finite. Let σ^2 represent the variance of $f(\mathcal{B}_1)$.

The algorithm we present in this section was introduced in [3, 23] for positive recurrent Markov chains with an accessible known atom. Similarly to the RBB, it consists on dividing the chain into $\mathcal{B}_1, \ldots, \mathcal{B}_{T(n)}$ regenerative blocks and then resampling blocks to form the empirical distribution of $\mathcal{B}_1, \ldots, \mathcal{B}_{T(n)}$. The main difference between the Regeneration-based bootstrap and the RBB is that in the former, the number of bootstrapped blocks is $T(n)$, hence, non-random conditionally to $\mathbf{X}^{(n)}$, while in the latter is random.

The full algorithm is as follows:

1. Count the number of visits $T_n(\boldsymbol{\alpha})$ to the atom $\boldsymbol{\alpha}$ up to time n, and divide the observed sample path $\mathbf{X}^{(n)} = (X_0, \ldots, X_n)$ into $T_n(\boldsymbol{\alpha}) + 1$ blocks, $\mathcal{B}_0$, $\mathcal{B}_1, \ldots, \mathcal{B}_{T_n(\boldsymbol{\alpha})-1}$, $\mathcal{B}^{(n)}_{T_n(\boldsymbol{\alpha})}$, corresponding to the pieces of the sample path between consecutive visits to the atom $\boldsymbol{\alpha}$. Drop the first and last (non-regenerative) blocks. Denote by $T(n)$ the number of remaining blocks.
2. Draw $T(n)$ bootstrap data blocks $\mathcal{B}^*_{1,T(n)}, ..., \mathcal{B}^*_{T(n),T(n)}$ independently from the empirical distribution $F_n = T(n)^{-1} \sum_{j=1}^{T(n)} \delta_{\mathcal{B}_j}$ of the blocks $\{\mathcal{B}_j\}_{1 \leqslant j \leqslant T(n)}$ conditioned on $\mathbf{X}^{(n)}$.
3. From the bootstrap data blocks generated at step 2, reconstruct a trajectory by binding the blocks together, getting the reconstructed sample path
$$X^{*(n)} = (\mathcal{B}^*_{1,T(n)}, ..., \mathcal{B}^*_{T(n),T(n)}).$$
Compute the statistic $G^*_n = G_n\left(X^{*(n)}\right)$.
4. If $Std_n = Std(\mathcal{B}_1, \ldots, \mathcal{B}_{T(n)})$ is an appropriate standardization of the original statistic G_n, compute $Std^*_n = Std(\mathcal{B}^*_{1,T(n)}, \ldots, \mathcal{B}^*_{T(n),T(n)})$.

As in the RBB case, the asymptotic result stated below shows the validity of this bootstrap scheme when used in estimations of integrals with respect to the invariant measure. In line with the conventions of bootstrap literature, we will write $Z^*_n \xrightarrow{d^*}_{a.s.} Z$ to denote the weak convergence almost surely along the data of the bootstrap random variables Z^*_n towards Z. This means that, for every $x \in \mathbb{R}$, $\mathbb{P}^*(Z^*_n \leqslant x) \xrightarrow{a.s.} \mathbb{P}(Z \leqslant x)$. For more details, see [16, pp. 2250].

Theorem 3.1 (Validity of the Regeneration based bootstrap)**.** *Under the same hypothesis of Theorem 2.1, we have*

$$\frac{\sqrt{T(n)}}{\widehat{\sigma}_{T(n)}} \left(\frac{1}{T(n)} \sum_{j=1}^{T(n)} \left(f(\mathcal{B}^*_{j,T(n)}) - \widehat{\mu}_{T(n)} \right) \right) \xrightarrow{d^*}_{a.s.} \boldsymbol{N}(0,1)\,.$$

Remark 3.1. In its original formulation for the positive recurrent case, the estimator used was $S_n(f)/n$, however, by Remark 1.2, it can not be used in the null recurrent case.

4. Simulations

In order to empirically compare the two bootstrap methodologies described in this paper, we devote this section to simulation examples. The code for all the experiments is available at https://github.com/carlosds731/boostrap_markov

As a model for the experiments, we will consider the simple symmetric random walk in $\mathbb{Z}$, that is

$$X_t = \begin{cases} 0 & , \quad t = 0 \\ \sum\limits_{k=1}^{t} Y_k & , \quad t \geqslant 1 \end{cases} \tag{17}$$

with $P(Y_i = 1) = P(Y_i = -1) = 1/2$. In this random walk, the state 0 is an atom and the invariant measure is $\pi_0(\{i\}) \equiv 1$ (see pp.1143 in [5]). Consider the function $f(k) = \frac{1}{k^2}$ if $k \neq 0$ and $f(0) = 0$, then

$$\int f(x)\, d\pi_0(x) = 2\sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{3}.$$

Our parameter of interest will be $\int f(x)\, d\pi_0(x)$ which we will estimate with $G_n = \sum_{i=1}^{T(n)} f(\mathcal{B}_i)/T(n)$. The bootstrap version of this statistic will be denoted by $G^*_{RBB,n}$ in the RBB case and by $G^*_{RgB,n}$ in the regeneration based scenario. Their standardized versions are defined as follows:

$$L_n := \frac{\sqrt{T(n)}}{\widehat{\sigma}_n}\left(G_n - \int f\, d\pi_{\boldsymbol{\alpha}}\right),$$

$$L^*_{RBB,n} := \frac{\sqrt{T^*(n, T(n))}}{\widehat{\sigma}_n}\left(G^*_{RBB,n} - G_n\right),$$

$$L^*_{RgB,n} := \frac{\sqrt{T(n)}}{\widehat{\sigma}_n}\left(G^*_{RgB,n} - G_n\right).$$

By Corollary 1.2 and Theorems 2.1 and 3.1 we have that $L_n \xrightarrow{d} \mathbf{N}(0,1)$, $L^*_{RBB,n} \xrightarrow{d^*}_p \mathbf{N}(0,1)$ and $L^*_{RgB,n} \xrightarrow{d^*}_{a.s.} \mathbf{N}(0,1)$.

4.1. Comparisons with the true distribution

To see the finite sample performance of both bootstrap methods, and compare its accuracy, for different values of n we have simulated a realization of the chain of length n and then applied both bootstrap methods 10^4 times, obtaining that

many samples of $L^*_{RBB,n}$ and $L^*_{RgB,n}$. We have then computed the empirical cumulative distribution function of these statistics and compared with the CDF of L_n.

The results of these simulations, presented in Figure 1, show that as n increases, the distributions of both $L^*_{RBB,n}$ and $L^*_{RgB,n}$ approximate the true distribution of L_n. Regarding the accuracy, the experiment gives empirical evidence, that, as in the positive recurrent case, the RBB provides a more accurate approximation of the true distribution than the regeneration based bootstrap or the asymptotic normal distribution.

Remark 4.1. For each n, the true distribution of L_n is unknown. To obtain a reliable approximation of its cumulative distribution function, we simulated 10^5 independent realizations of $\mathbf{X}$ of length n. We then used these samples to compute the empirical cumulative distribution function of L_n.

4.2. Coverage probability

The bootstrap methods' first-order correctness established in this paper allows us to use the quantiles of $L^*_{RBB,n}$ and $L^*_{RgB,n}$ to construct confidence intervals for $\int f\, d\pi_{\boldsymbol{\alpha}}$. Let $q^*_{RBB}(\alpha)$ and $q^*_{RgB}(\alpha)$ represent the α-quantiles of $L^*_{RBB,n}$ and $L^*_{RgB,n}$ respectively. The bootstrap confidence intervals are then given by:

$$I^*_{RBB,n} = \left[G_n - \frac{\widehat{\sigma}_n}{\sqrt{T(n)}} q^*_{RBB}(1-\alpha/2), G_n - \frac{\widehat{\sigma}_n}{\sqrt{T(n)}} q^*_{RBB}(\alpha/2) \right],$$

$$I^*_{RgB,n} = \left[G_n - \frac{\widehat{\sigma}_n}{\sqrt{T(n)}} q^*_{RgB}(1-\alpha/2), G_n - \frac{\widehat{\sigma}_n}{\sqrt{T(n)}} q^*_{RgB}(\alpha/2) \right].$$

Figure 2a shows the coverage probabilities of $I^*_{RBB,n}$ and $I^*_{RgB,n}$ for $\alpha = 0.05$ for different values of n while figure 2b shows the average length of these confidence intervals. For comparisons, we have also included in figures 2a and 2b the coverage probabilities and average interval length of the confidence intervals obtained when we use the normal approximation.

As expected, as n gets larger, the coverage probability of the confidence intervals approaches the desired level (0.95), while the average length decreases. This experiment reinforces the idea that the RBB provides a better approximation than the regeneration based bootstrap, as it produces confidence intervals with higher coverage probability and very similar length. In comparison with the asymptotic distribution, the RBB generates confidence intervals with narrower lengths and similar coverage probabilities. This could be explained by noticing that the asymptotic distribution, being symmetric, does not take into account the possible asymmetry of the underlying distribution.

Remark 4.2. To determine the coverage probability for a given n, we simulated 10^4 independent realizations of $\mathbf{X}$ with length n. We then applied both bootstrap methods 10^4 times, generating the same number of samples for $L^*_{RBB,n}$ and $L^*_{RgB,n}$, which were subsequently used to compute their bootstrap quantiles.

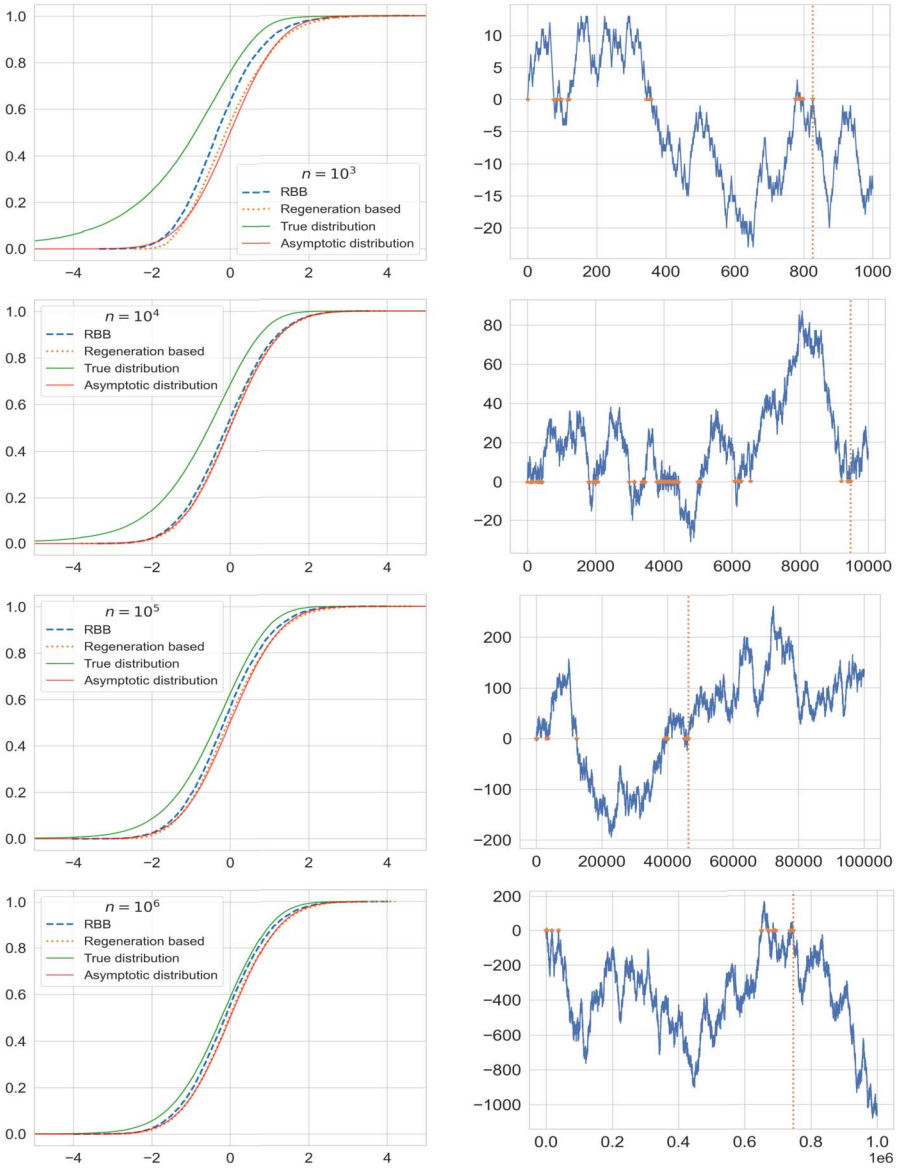
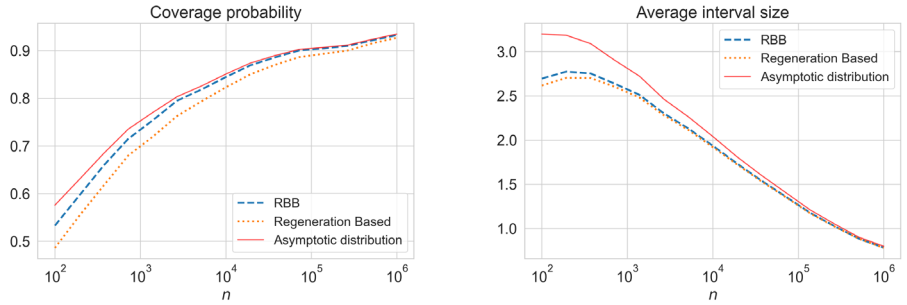


FIG 1. *Distribution functions of* $L^*_{RBB,n}$ *and* $L^*_{RgB,n}$*, the true distribution of* L_n *and the asymptotic distribution (standard normal) for different values of* n *(left column) and the realization of* $\boldsymbol{X}$ *from where the samples of* $L^*_{RBB,n}$ *and* $L^*_{RgB,n}$ *were obtained (right column), the orange stars mark the regeneration times, while the orange dotted lines indicate the end of the last complete block.*



(a) Coverage probabilities (b) Average interval length

FIG 2. *Coverage probabilities and average interval length of the* $I^*_{RBB,n}$ *and* $I^*_{RgB,n}$ *at* 95% *confidence level. The x-axis is in logarithmic scale.*

5. Conclusions and perspectives

In this work we have proved the first order validity of the Regenerative block-bootstrap and Regeneration based bootstrap for atomic β-null recurrent Markov chains. Up to our knowledge these are the first bootstrap method whose validity has been established for these type of non-stationary Markov chains.

In terms of extending the methods to non-atomic chains, it is possible to apply the Nummelin splitting technique [52, 53], following the approach described in [11, Section 3]. This construction involves "extending" the chain to make it atomic, then applying the bootstrap to the extended chain. However, to establish the validity of these bootstrap procedures for non-atomic β-null recurrent chains, several additional steps are required. First, we need to derive a uniform rate of convergence on a small set for the transition kernel estimator (which has been done at specific points by [37] but not uniformly). We also require new exponential inequalities to obtain rate of convergence. Finally, it must be shown that the same type of coupling used in [11] still holds in this case. This will be the subject of further investigations.

6. Proofs

6.1. *Proof of Lemma 1.2*

For the proof of Lemma 1.2 we need the following result, which appears as part A.3 of Theorem A.1 in [37].

Lemma 6.1. *Let* A_n *and* B_n *be a pair of stochastic processes which are càdlàg, where* A_n *is non-negative and non-decreasing. Let* B *denote a Brownian motion defined for* $t \in \mathbb{R}$ *and let* A *denote a strictly increasing non-negative process with independent increments,* $A(0) \equiv 0$ *and with no fixed jumps. Assume that*

$B_n \xrightarrow{\mathscr{D}_{[0,+\infty)}} B$ *and* $A_n \xrightarrow{\mathscr{D}_{[0,+\infty)}} A$. *Then,* $A_n^{(-1)} \xrightarrow{\mathscr{D}_{[0,+\infty)}} A^{(-1)}$ *and*

$$\left(A_n^{(-1)}(t), \frac{B_n \circ A_n^{(-1)}(t)}{\sqrt{A_n^{(-1)}(t)}} \right) \xrightarrow{d} (A^{(-1)}(t), Z) \quad \forall t \in (0,1],$$

where Z *is standard normal random variable independent of* $A^{(-1)}(t)$.

To prove Lemma 1.2, let $W_k = \sigma^{-1}(X_k - \mu)$, then $\{W_k\}_{k=1}^{\infty}$ is an i.i.d. sequence with $\mathbb{E}W_k = 0$ and $\operatorname{Var} W_k = 1$ for all k. Define the following continuous time process for $t \geqslant 0$

$$Q_n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} W_k. \tag{18}$$

By Theorem 23 and Example 24 in [63], $Q_n \xrightarrow{\mathscr{D}_{[0,+\infty)}} B$ and given that u_n is an unbounded increasing sequence, we also have that Q_{u_n} converges weakly to B in $\mathscr{D}_{[0,+\infty)}$.

The conditions imposed to the process N_n allow us to apply Lemma 6.1 with $A_n = S_n$ and $B_n = Q_{u_n}$. Taking into account that N_n is equivalent to $S_n^{(-1)}$ we obtain that for all $t > 0$

$$\frac{Q_{u_n}(N_n(t))}{\sqrt{N_n(t)}} \xrightarrow{d} \mathbf{N}(0,1). \tag{19}$$

Using that $N(\lfloor nt \rfloor) = u_n N_n(t)$, we get

$$Q_{u_n}(N_n(t)) = \frac{\sigma^{-1}}{\sqrt{u_n}} \sum_{j=1}^{N(\lfloor nt \rfloor)} (X_j - \mu), \tag{20}$$

and Lemma 1.2 follows after plugging (20) into (19) and taking $t = 1$.

6.2. *Proof of Corollary 1.1*

We assume, at first, that θ is bounded, that is, there exists a constant K such that $0 < \theta < K$ with probability 1. Without loss of generality, assume the u_n are integers. Define the process

$$N_n(t) = \begin{cases} \frac{tN(n)}{u_n} & , \text{ if } \frac{N(n)}{u_n} < 1 \\ t\theta & , \text{ otherwise} \end{cases}.$$

As stated in pp. 147 of [13], this process converges to the process $t\theta$ and trivially satisfies the conditions of Lemma 1.2 (using $S_n(t) = \frac{t}{\theta}$, $S_n^{-1}(t) = t\theta$).

The case when K is unbounded can be treated by following the same argument as in pp. 148 of [13].

6.3. Proof of Theorem 1.3

Recall from Section 1.3 that, by the Strong Markov Property, the sequence $\{f(\mathcal{B}_j)\}_{j=1}^{+\infty}$ is i.i.d. with mean $\int f d\pi_{\boldsymbol{\alpha}}$ and variance σ^2. Consider the processes $T_n(t)$ and C_n defined in (8)

$$T_n(t) = \frac{T(\lfloor nt \rfloor)}{u(n)} \quad , \quad C_n(t) = \frac{1}{v(n)} \sum_{k=0}^{\lfloor nt \rfloor} \ell(\mathcal{B}_k).$$

By Theorem 1.1, we can apply Lemma 1.2 with $X_i = f(\mathcal{B}_i)$, $\mu = \int f d\pi_{\boldsymbol{\alpha}}$, $N(n) = T(n)$ and $u_n = n^\beta L(n)$, which completes the proof of (15). In order to prove (16), denote by W_n the left-hand side of (15), then

$$\frac{\sqrt{T(n)}}{\sigma} \left(\frac{S_n(f)}{T(n)} - \int f \, d\pi_{\boldsymbol{\alpha}} \right) = W_n + \frac{f(\mathcal{B}_0)}{\sqrt{T(n)}} + \frac{f(\mathcal{B}_{(n)})}{\sqrt{T(n)}},$$

therefore, (16) will follow from (15) if we show that both $f(\mathcal{B}_0)/\sqrt{T(n)}$ and $f(\mathcal{B}_{(n)})/\sqrt{T(n)}$ converge to 0 at least in probability.

The random variable $f(\mathcal{B}_0)$ is almost surely bounded and $T(n)$ converges almost surely to $+\infty$, therefore $f(\mathcal{B}_0)/\sqrt{T(n)}$ converges to 0 almost surely. For the other term, first notice that for all $n \in \mathbb{N}$ we have

$$\frac{|f(\mathcal{B}_{(n)})|}{\sqrt{T(n)}} \leqslant \frac{|f|(\mathcal{B}_{(n)})}{\sqrt{T(n)}} \leqslant \frac{|f|(\mathcal{B}_{T(n)+1})}{\sqrt{T(n)}}.$$

The random variables $\{|f|(\mathcal{B}_j)\}_{j \geqslant 1}$ are i.i.d. with finite second moment, therefore, by Lemma 1 in [6], $|f|(\mathcal{B}_n)/\sqrt{n}$ converges to 0 a.s. Since $T(n)$ converges to $+\infty$ almost surely, Theorem 6.8.1 in [30] implies that $|f|(\mathcal{B}_{T(n)+1})/\sqrt{T(n)}$ converges to 0 with probability 1, which concludes the proof of Theorem 1.3.

6.4. Proof of Theorem 2.1

Assume we have observed the chain until time n, i.e., $\mathbf{X}^{(n)} = X_0, X_1, \ldots, X_n$, and we have extracted the $T(n)$ regeneration blocks: $\mathcal{B}_1, \ldots, \mathcal{B}_{T(n)}$.

Now we start to sequentially bootstrap data blocks $\mathcal{B}^*_{1,T(n)}, \ldots, \mathcal{B}^*_{k,T(n)}$ independently from the empirical distribution $F_{T(n)} = T(n)^{-1} \sum_{j=1}^{T(n)} \delta_{\mathcal{B}_j}$ of the blocks $\{\mathcal{B}_j\}_{1 \leqslant j \leqslant T(n)}$, conditioned on $\mathbf{X}^{(n)}$, until the length of the bootstrap data series, $\ell^*(k) = \sum_{j=1}^{k} \ell\left(\mathcal{B}^*_{1,T(n)}\right)$, is larger than n.

For each m, define

$$T^*(m, T(n)) = \max \left\{ k : \sum_{j=1}^{k} \ell\left(f(\mathcal{B}^*_{j,T(n)})\right) \leqslant m \right\}, \tag{21}$$

$$U^*\left(m, T\left(n\right)\right) = \frac{\sqrt{m}}{\widehat{\sigma}_{T(n)}}\left(\frac{1}{m}\sum_{j=1}^{m}\left(f(\mathcal{B}^*_{j,T(n)}) - \mu_{T(n)}\right)\right). \tag{22}$$

Theorem 2.1 will be proved if we show that, for all $x \in \mathbb{R}$ it holds that

$$\mathbb{P}^*\left(U^*\left(T^*(n, T\left(n\right)), T\left(n\right)\right) \leqslant x\right) \xrightarrow{p} \Phi(x), \tag{23}$$

where Φ is the cumulative distribution function of a standard normal random variable and $\mathbb{P}^*(\bullet) = \mathbb{P}(\bullet \mid \mathbf{X}^{(n)})$ denotes the conditional probability given $\mathbf{X}^{(n)}$.

Given that we will bootstrap $T^*(n, T\left(n\right))$ terms, which is a random quantity conditionally to the data, we will use Lemma 6.1 to prove (23). In order to do this we need, conditionally to the data:

1. Find a process $S^*_{n,T(n)}(t)$ that is non-negative, non-decreasing that converges in $\mathscr{D}_{[0,+\infty)}$ to a process S^* that is non-negative, strictly increasing, has independent increments, no fixed jumps and $S^*\left(0\right) \equiv 0$.
2. Show that $T^*_{n,T(n)}(t) = T^*\left(\lfloor nt \rfloor\right)/T\left(n\right) = T^*\left(\lfloor nt \rfloor, T\left(n\right)\right)/T\left(n\right)$ is equivalent in $\mathscr{D}_{[0,+\infty)}$ to $S^{*(-1)}_{n,T(n)}$.
3. Find a process $Q^*_{n,T(n)}(t)$ that converges in $\mathscr{D}_{[0,+\infty)}$ to a Brownian motion when n goes to $+\infty$. This process should satisfy, for some $t > 0$

$$U^*\left(T^*(n, T\left(n\right)), T\left(n\right)\right) = \frac{Q^*_{n,T(n)} \circ T^*_{n,T(n)}(t)}{\sqrt{T^*_{n,T(n)}(t)}}. \tag{24}$$

A natural choice for $Q^*_{n,T(n)}$, which satisfies (24) for $t = 1$, is

$$Q^*_{n,T(n)}(t) = \frac{\sqrt{T\left(n\right)}}{\widehat{\sigma}_{T(n)}}\left(\frac{1}{T\left(n\right)}\sum_{j=1}^{\lfloor T(n)t \rfloor}\left(f(\mathcal{B}^*_{j,T(n)}) - \mu_{T(n)}\right)\right). \tag{25}$$

Take $S^*_{n,T(n)}(t)$ as

$$S^*_n(t) = \frac{1}{v^*(T\left(n\right))}\sum_{i=1}^{\lfloor T(n)t \rfloor}\ell\left(\mathcal{B}^*_{i,T(n)}\right), \tag{26}$$

where $v^*(T\left(n\right)) = \sum_{i=0}^{T(n)}\ell(\mathcal{B}_i)$.

Following the notation of [39], let $Y_i = \ell(\mathcal{B}_i)$ and let $Y_{1,n} \geqslant Y_{2,n} \geqslant \cdots \geqslant Y_{n,n}$ be the order statistics of the sizes of the first n blocks, and take $Z_{k,n} = Y_{k,n}/v(n)$ where $v(n)$ is as in (7). By Theorem 1 in [39],

$$Z^{(n)} = (Z_{1,n}, Z_{2,n}, \ldots, Z_{n,n}, 0, 0, \ldots) \xrightarrow{d} (Z_1, Z_2, \ldots) = Z, \tag{27}$$

where $Z_k = \left(E_1 + \cdots + E_k\right)^{-\frac{1}{\beta}}$ and E_i is a sequence of i.i.d. of exponential random variables with mean 1. By Skorokhod-Dudley-Wichura Theorem (see

pp. 1171 in [39] and pp. 476 in [7]) we can choose a probability space such that, without changing the distribution of the left-hand side of (27),

$$Z^{(n)} \xrightarrow{a.s.} Z. \tag{28}$$

The following Lemma shows that in that space, conditionally to the data, the process $S^*_{n,T(n)}$ converges in $\mathscr{D}_{[0,+\infty)}$.

Lemma 6.2. *Suppose that (28) holds. Let $\lambda^*_j(t)$ be independent Poisson processes with parameter 1, and K a positive constant. Define*

$$R^*(t) = \sum_{j=1}^{+\infty} Z_j \left(\lambda^*_j(t) - t\right) \quad \text{and} \quad S^*(t) = KR^*(t) + t.$$

Then, $T(n)/u(n)$ converges almost surely to a positive random variable and

$$S^*_{n,T(n)} \xrightarrow{\mathscr{D}_{[0,+\infty)}} S^* \quad \text{and} \quad S^{*(-1)}_{n,T(n)} \xrightarrow{\mathscr{D}_{[0,+\infty)}} S^{*(-1)} \tag{29}$$

almost surely along the data. Moreover, the process S^* is non-negative, strictly increasing, continuous, with independent increments and $S^*(0) \equiv 0$.

Proof. When (28) holds, by Theorem 1 and Remark 1.3 in [47],

$$\frac{1}{v(n)} \sum_{j=1}^{n} \ell(\mathcal{B}_j) \xrightarrow{a.s.} \sum_{j=1}^{+\infty} Z_j.$$

The length of the first block, $\ell(\mathcal{B}_0)$, is finite with probability 1 and does not depend on n, hence $\ell(\mathcal{B}_0)/v(n)$ converges almost surely to 0. This implies that

$$\frac{1}{v(n)} \sum_{j=0}^{n} \ell(\mathcal{B}_j) \xrightarrow{a.s.} \sum_{j=1}^{+\infty} Z_j. \tag{30}$$

In (7), we defined $v(z)$ as the inverse of $u(z) = z^\beta L(z)$, then, by Proposition 1.5.15 in [14], $v(z) \sim z^{1/\beta} L_1(z)$ where L_1 is a slowly varying function, hence, $v(n)/v(\lfloor nt \rfloor) \to t^{-1/\beta}$, and we have that

$$\frac{1}{v(n)} \sum_{j=0}^{\lfloor nt \rfloor} \ell(\mathcal{B}_j) \xrightarrow{a.s.} t^{\frac{1}{\beta}} \sum_{j=1}^{+\infty} Z_j \quad \forall t > 0. \tag{31}$$

For each $t > 0$, let

$$S_n(t) = \frac{1}{v(n)} \sum_{j=0}^{\lfloor nt \rfloor} \ell(\mathcal{B}_j)\ ,\ S^{(-1)}_n(t) = \inf\{x > 0 : S_n(x) > t\}\ ,\ S(t) = t^{1/\beta} \sum_{j=1}^{+\infty} Z_j,$$

and define the three processes as 0 on $t = 0$. By (31) and the Continuous Mapping Theorem, $S^{(-1)}_n \xrightarrow{a.s.} S^{-1}$.

Similar to what is described on page 1141 in [5], suppose that y is such that $y < S_n^{(-1)}(1)$. Then, since $S_n(y) < 1$, it follows that $\sum_{j=0}^{\lfloor ny \rfloor} \ell(\mathcal{B}_j) < v(n)$. Consequently, we have $T(\lfloor v(n) \rfloor) \geqslant \lfloor ny \rfloor > ny - 1$. This in turn implies that $T(\lfloor v(n) \rfloor)/n \geqslant y - 1/n \geqslant S_n^{(-1)}(1) - 1/n$ for all n. Similarly, but taking $y > S_n^{(-1)}$, we show that $T(\lfloor v(n) \rfloor)/n \leqslant S_n^{(-1)}(1) + 1/n$ for all n. Then,

$$S_{u(n)}^{(-1)}(1) - \frac{1}{u(n)} \leqslant \frac{T(\lfloor v(u(n)) \rfloor)}{u(n)} \leqslant S_{u(n)}^{(-1)}(1) + \frac{1}{u(n)}. \tag{32}$$

The first part of the lemma now follows from (32), the convergence of $S_{u(n)}^{(-1)}(1)$ to $S^{-1}(1)$ and the fact that $u(v(n)) = n$ for n big enough.

To show (29), consider the following process, which was studied in [7],

$$Z_{m,n}^*(t) = \frac{1}{v(n)} \sum_{j=1}^{\lfloor mt \rfloor} \left(\ell(\mathcal{B}_{j,n}^*) - \frac{1}{n} \sum_{i=1}^{n} \ell(\mathcal{B}_i) \right).$$

By Corollary 1.2 in [7] (and its proof[4]), we see that when (28) holds, for any m_n such that $m_n/n \to c$, conditionally to the data, the process $Z_{m_n,n}^*$ converges weakly in $\mathscr{D}([0,1])$ to $R^*(ct)$. Let $C > 1$, on $[0, C]$ define the process

$$W_n^*(t) = \frac{1}{v(n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(\ell(\mathcal{B}_{j,n}^*) - \frac{1}{n} \sum_{i=1}^{n} \ell(\mathcal{B}_i) \right).$$

Notice that $W_n^*(t) = Z_{nC,n}^*(t/C)$, hence, $W_n^* \xrightarrow{\mathscr{D}_{[0,C]}} R^*$ as $n \to +\infty$. Because this convergence holds for arbitrary $C > 0$, by Lemma 1.3.ii in [38] we have that $W_n^* \xrightarrow{\mathscr{D}_{[0,+\infty)}} R^*$, and therefore, $W_{T(n)}^* \xrightarrow{\mathscr{D}_{[0,+\infty)}} R^*$.

The process $S_{n,T(n)}^*$ can be written as

$$S_{n,T(n)}^*(t) = \frac{v(T(n))}{v^*(T(n))} W_{T(n),T(n)}^*(t) + \frac{\lfloor T(n)\, t \rfloor}{T(n)}. \tag{33}$$

Notice that

$$\frac{v(T(n))}{v^*(T(n))} = \left(\frac{1}{v(T(n))} \sum_{j=0}^{T(n)} \ell(\mathcal{B}_j) \right)^{-1},$$

then, conditionally to the data, it converges to a positive constant K by equation (30). Equation (29) now follows from the convergence of $W_{T(n),T(n)}^*$ and (33). The continuity of S^* was shown in pp. 466 of [7], and the rest of the properties are directly deduced from the form of R^*. □

The next Lemma handles the equivalence of $T_{n,T(n)}^*$ and $S_{n,T(n)}^{*(-1)}$ in $\mathscr{D}_{[0,+\infty)}$.

[4] In [7], they standardize by $T_n = \max_{1 \leqslant k \leqslant n} l(\mathcal{B}_k)$ but from the proof is clear that the result remains valid if we standardize by $v(n)$ (b_n in their notation).

Lemma 6.3. *Under the same hypothesis of Lemma 6.2, the processes* $T^*_{n,T(n)}$ *and* $S^{*(-1)}_{n,T(n)}$ *are equivalent in* $\mathscr{D}_{[0,+\infty)}$.

Proof. The proof of this result follows the proof of Theorem 3.2 on [37] with slight modifications.

We need to show that, for any $\varepsilon > 0$ given,

$$\mathbb{P}\left(\sup_{0<t\leqslant K}\left|T^*_{n,T(n)}(t) - S^{*(-1)}_{n,T(n)}(t)\right| > \varepsilon\right) \to 0 \quad \forall K > 0. \tag{34}$$

To prove this, we will show that

$$\mathbb{P}\left(\sup_{0<t\leqslant K}\left|T^*_{v^*(T(n)),T(n)}(t) - S^{*(-1)}_{n,T(n)}(t)\right| > \varepsilon\right) \to 0 \quad \forall K > 0, \tag{35}$$

$$\mathbb{P}\left(\sup_{0<t<K}\left|T^*_{v^*(T(n)),T(n)}(t) - T^*_{n,T(n)}(t)\right| > \varepsilon\right) \to 0 \quad \forall K > 0. \tag{36}$$

from where (34) will follow by triangular inequality.

Let $\eta > 0$

$$\begin{aligned}\left\{S^{*(-1)}_{n,T(n)}(t) < \eta\right\} &\subseteq \left\{S^*_{n,T(n)}(\eta) > t\right\} \\ &= \left\{\frac{1}{v^*(T(n))}\sum_{i=1}^{\lfloor T(n)\eta\rfloor}\ell\left(\mathcal{B}^*_{i,T(n)}\right) > t\right\} \\ &= \left\{\sum_{i=1}^{\lfloor T(n)\eta\rfloor}\ell\left(\mathcal{B}^*_{i,T(n)}\right) > tv^*(T(n))\right\} \\ &= \left\{\frac{T^*(\lfloor v^*(T(n))t\rfloor, T(n))}{T(n)} < \frac{\lfloor T(n)\eta\rfloor}{T(n)}\right\}. \end{aligned} \tag{37}$$

Because $T(v^*(n)) = n$, we can write,

$$T^*_{v^*(T(n)),T(n)}(t) = \frac{T^*(\lfloor v^*(T(n))\,t\rfloor, T(n))}{u^*(v^*(T(n)))} = \frac{T^*(\lfloor v^*(T(n))\,t\rfloor, T(n))}{T(n)},$$

therefore, equation (37) becomes

$$\left\{S^{*(-1)}_{n,T(n)}(t) < \eta\right\} \subseteq \left\{T^*_{v^*(T(n)),T(n)}(t) < \frac{\lfloor T(n)\eta\rfloor}{T(n)}\right\}. \tag{38}$$

Similarly, we obtain that

$$\left\{S^{*(-1)}_{n,T(n)}(t) > \eta\right\} \subseteq \left\{T^*_{v^*(T(n)),T(n)}(t) \geqslant \frac{\lfloor T(n)\eta\rfloor}{T(n)}\right\}. \tag{39}$$

Let $\varepsilon_1 \in (0,1)$ be fixed and take $\eta_1 < \eta_2$, then, by (39) and (38),

$$\left\{\eta_1 \leqslant S^{*(-1)}_{n,T(n)}(t) < \eta_2\right\} \subseteq \left\{\eta_1(1-\varepsilon_1) < S^{*(-1)}_{n,T(n)}(t) < \eta_2\right\}$$

$$\subseteq \left\{ \frac{\lfloor T(n)\eta_1(1-\varepsilon_1)\rfloor}{T(n)} \leqslant T^*_{v^*(T(n)),T(n)}(t) < \frac{\lfloor T(n)\eta_2\rfloor}{T(n)} \right\}.$$

This means, that, if $S^{*(-1)}_{n,T(n)}(t) \in [\eta_1, \eta_2)$, then

$$\frac{\lfloor T(n)\eta_1(1-\varepsilon_1)\rfloor}{T(n)} - \eta_2 < T^*_{v^*(T(n)),T(n)}(t) - S^{*(-1)}_{n,T(n)}(t) < \frac{\lfloor T(n)\eta_2\rfloor}{T(n)} - \eta_1,$$

which implies that, if $S^{*(-1)}_{n,T(n)}(t) \in [\eta_1, \eta_2)$, then

$$\left| T^*_{v^*(T(n)),T(n)}(t) - S^{*(-1)}_{n,T(n)}(t) \right| \leqslant \eta_2 - \eta_1 + \varepsilon_1\eta_1 + \frac{1}{T(n)}. \tag{40}$$

Let $\varepsilon > 0$ be fixed. For any s we have

$$\mathbb{P}\left(\sup_{t\leqslant K} \left|\xi^*_{n,T(n)}(t)\right| > \varepsilon \right) \leqslant \mathbb{P}\left(\sup_{t\leqslant K} \left|\xi^*_{n,T(n)}(t)\right| > \varepsilon, \sup_{t\leqslant K} S^{*(-1)}_{n,T(n)}(t) < s \right) + \mathbb{P}\left(\sup_{t\leqslant K} S^{*(-1)}_{n,T(n)}(t) \geqslant s \right),$$

where $\xi^*_{n,T(n)}(t) = T^*_{v^*(T(n)),T(n)}(t) - S^{*(-1)}_{n,T(n)}(t)$.

By (29),

$$\lim_{s\uparrow\infty} \lim_{n\to\infty} \mathbb{P}\left(\sup_{t\leqslant K} S^{*(-1)}_{n,T(n)}(t) \geqslant s \right) = 0.$$

Therefore, for any $\delta > 0$ we can choose s_0 such that, for n big enough,

$$\mathbb{P}\left(\sup_{t\leqslant K} S^{*(-1)}_{n,T(n)}(t) \geqslant s_0 \right) < \delta.$$

By (40), $\sup_{t\leqslant K} S^{*(-1)}_{n,T(n)}(t) < s_0$ implies that

$$\left|\xi^*_{n,T(n)}(t)\right| \leqslant \eta_2 - \eta_1 + \varepsilon_1\eta_1 + \frac{1}{T(n)} \quad \forall t \in [0,K] \ , \ \forall \varepsilon_1 \in (0,1).$$

Choose $\eta_0, \ldots, \eta_L, N_1, \varepsilon_1$ with $\eta_0 = 0 < \eta_1 < \ldots < \eta_{L-1} < \eta_L = s_0$ such that $\eta_i - \eta_{i+1} < \varepsilon/3$ for all i. Let $\varepsilon_1 < \varepsilon/s_0$ and choose N_1 such that $1/T(N_1) < \varepsilon/3$.

Notice that for all $t \in [0,K]$ there is only one $i_{n,t}$ such that $S^{*(-1)}_{n,T(n)}(t)$ belongs to $\left[\eta_{i_{n,t}}, \eta_{i_{n,t}+1}\right)$, then, by (40)

$$\left|\xi^*_{n,T(n)}(t)\right| \leqslant \eta_{i_{n,t}} - \eta_{i_{n,t}+1} + \varepsilon_1\eta_1 + \frac{1}{T(n)} \leqslant \varepsilon \quad \forall t \in [0,K] \ , \ \forall n > N_1,$$

whenever $S^{*(-1)}_{n,T(n)}(t) < s_0$. This implies that

$$\mathbb{P}\left(\sup_{t\leqslant K} \left|\xi^*_{n,T(n)}(t)\right| > \varepsilon, \sup_{t\leqslant K} S^{*(-1)}_{n,T(n)}(t) < s_0 \right) = 0 \quad \forall n \geqslant N_1.$$

Hence,

$$\mathbb{P}\left(\sup_{t\leqslant K}\left|\xi^*_{n,T(n)}(t)\right|>\varepsilon\right)<\delta \quad \forall n>N_1, \tag{41}$$

which implies (35).

Now we turn to the proof of (36).

According to the definition of v^*, $v^*(T(n))=\sum_{i=0}^{T(n)}\ell(\mathcal{B}_i)\leqslant n$, therefore,

$$T^*_{v^*(T(n)),T(n)}(t)=\frac{T^*(\lfloor v^*(T(n))\,t\rfloor,T(n))}{T(n)}\leqslant\frac{T^*(\lfloor nt\rfloor,T(n))}{T(n)}$$
$$\leqslant T^*_{n,T(n)}(t)\quad\forall n,t.$$

Notice that $v^*(T(n)+1)=\sum_{i=0}^{T(n)+1}\ell(\mathcal{B}_i)>n$, therefore,

$$T^*_{n,T(n)}(t)\leqslant T^*_{v^*(T(n)+1),T(n)}(t)\frac{T(n)+1}{T(n)}\quad\forall n,\ t.$$

Hence,

$$T^*_{v^*(T(n)),T(n)}(t)\leqslant T^*_n(t)\leqslant T^*_{v^*(T(n)+1)}(t)\frac{T(n)+1}{T(n)}\quad\forall n,t.$$

Equation (36) now follows from the convergence of both $T^*_{v^*(T(n)),T(n)}$ and $T^*_{v^*(T(n)+1),T(n)}$ to $S^{*(-1)}$ and the fact that $(T(n)+1)/T(n)$ converges almost surely to 1. □

By (25), Lemmas 6.1, 6.2 and 6.3 we have that, in a space where (28) holds, the convergence in (23) holds almost surely. Therefore, in the original space we have the weakly-weakly convergence[5]

$$\forall x\in\mathbb{R}\quad\mathbb{P}^*\left(U^*\left(T^*(n,T(n)),T(n)\right)\leqslant x\right)\xrightarrow{d}\Phi(x). \tag{42}$$

However, given that the right-hand side of (42) is a constant for each x, the convergence in (42) can be improved to convergence in probability, which completes the proof.

6.5. *Proof of Theorem 3.1*

This proof follows the line of the proof of Theorem 2.1 in [12]. As in that paper, let Γ_2 be the set of distribution functions G satisfying $\int x^2dG(x)<\infty$ and

[5]The weakly-weakly convergence, introduced in [16] is the translation of the concept of weak convergence of random measures to the probabilistic setting, that is, for random variables (Z,X) and (Z_n,X_n) defined on possibly different probability spaces, the weakly-weakly convergence of $Z_n|X_n$ to $Z|X$ is defined by the fact $\mathbb{E}[g(Z_n)|X_n]\xrightarrow{d}\mathbb{E}[g(Z)|X]$ for all bounded and continuous functions g. For a detailed description of this concept as well as other examples of its application in the bootstrap setting, please refer to pp. 2550 and Appendix A in [16].

define the following notion of convergence in Γ_2

$$G_n \Rightarrow G \quad \text{iff} \quad G_n \to G \text{ weakly and } \int x^2 dG_n(x) \to \int x^2 dG(x). \tag{43}$$

Denote by d_2 a Mallows metric that metricizes the $\Rightarrow$ convergence in Γ_2 (see details in Section 8 of [12])

If $Y_1, \ldots, Y_n$ are i.i.d. random variables with common distribution G, denote by $G^{(m)}$ the distribution of

$$m^{-\frac{1}{2}} \sum_{j=1}^{m} (Y_j - \mathbb{E}Y_j).$$

By pp. 1198 in [12], if $G, H \in \Gamma_2$ then $G^{(m)}$ and $H^{(m)}$ are also in Γ_2 and

$$d_2\left(G^{(m)}, H^{(m)}\right) \leqslant d_2(G, H). \tag{44}$$

Let F be the distribution of $f(\mathcal{B}_1)$ and denote by F_n the empirical distribution function of $f(\mathcal{B}_1), \ldots, f(\mathcal{B}_n)$. By (2.1) in [12] and the fact that $T(n) \to +\infty$ a.s., $F_{T(n)} \Rightarrow F$ along almost almost all sample paths, hence, conditionally to the data

$$d_2\left(F_{T(n)}, F\right) \to 0. \tag{45}$$

Denote by N_σ a standard distribution with mean 0 and variance σ^2. By Proposition 1.3,

$$d_2\left(F^{(T(n))}, N_\sigma\right) \to 0. \tag{46}$$

Conditionally to the data, the distribution of

$$\sqrt{T(n)} \left(\frac{\sum_{j=1}^{T(n)} \left(f\left(\mathcal{B}^*_{j,T(n)}\right) - \frac{1}{T(n)} \sum_{i=1}^{T(n)} f(\mathcal{B}_i) \right)}{T(n)} \right)$$

is $F^{(T(n))}_{T(n)}$, then, conditionally to the data,

$$d_2\left(F^{(T(n))}_{T(n)}, N_\sigma\right) \leq d_2\left(F^{(T(n))}_{T(n)}, F^{(T(n))}\right) + d_2\left(F^{(T(n))}, N_\sigma\right)$$

which goes to 0 by (45) and (46). The theorem now follows by (43), (13) and Slutsky's theorem.

Acknowledgments

The author would like to thank the two anonymous referees and the Associate Editor for their constructive comments that greatly improved the quality of this paper. A significant part of this work was conducted during the author's PhD at Université Pars-Nanterre.

Funding

This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01) and has also been funded by a post doctoral grant of the Hi! Paris foundation.

References

[1] Alexander, K. (2011). Excursions and Local Limit Theorems for Bessel-like Random Walks. *Electronic Journal of Probability* **16**. https://doi.org/10.1214/ejp.v16-848. MR2749771

[2] Anscombe, F. J. (1952). Large-sample theory of sequential estimation. *Mathematical Proceedings of the Cambridge Philosophical Society* **48**. https://doi.org/10.1017/s0305004100076386. MR0051486

[3] Athreya, K. B. and Fuh, C. D. (1992). Bootstrapping Markov chains: countable case. *Journal of Statistical Planning and Inference* **33** 311-331. https://doi.org/10.1016/0378-3758(92)90002-A. MR1200651

[4] Athreya, K. B. (1987). Bootstrap of the Mean in the Infinite Variance Case. *The Annals of Statistics* **15** 724 – 731. https://doi.org/10.1214/aos/1176350371. MR0888436

[5] Athreya, K. B. and Roy, V. (2015). Estimation of integrals with respect to infinite measures using regenerative sequences. *Journal of Applied Probability* **52** 1133 – 1145. https://doi.org/10.1239/jap/1450802757. MR3439176

[6] Athreya, K. B. and Roy, V. (2016). General Glivenko-Cantelli Theorems. *Stat* **5**. https://doi.org/10.1002/sta4.128. MR3589269

[7] Berkes, I., Horváth, L. and Schauer, J. (2010). Non-central limit theorems for random selections. *Probability theory and related fields* **147** 449–479. MR2639712

[8] Bertail, P. (1994). Un test Bootstrap dans un modéle AR(1). *Annals of Economics and Statistics* **36** 57-79. MR1462741

[9] Bertail, P. and Clémençon, S. (2004). Edgeworth expansions of suitably normalized sample mean statistics for atomic Markov chains. *Probability Theory and Related Fields* **130** 388–414. https://doi.org/10.1007/s00440-004-0360-0. MR2095936

[10] Bertail, P. and Clémençon, S. (2007). Second-order properties of regeneration-based bootstrap for atomic Markov chains. *TEST* **16** 109–122. https://doi.org/10.1007/s11749-006-0004-z. MR2313804

[11] Bertail, P. and Clémençon, S. (2006). Regenerative block bootstrap for Markov chains. *Bernoulli* **12** 689–712. MR2248233

[12] Bickel, P. J. and Freedman, D. A. (1981). Some Asymptotic Theory for the Bootstrap. *The Annals of Statistics* **9**. https://doi.org/10.2307/2240410. MR0630103

[13] Billingsley, P. (1968). *Convergence of probability measures*. Wiley. MR0233396

[14] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular variation. Encyclopedia of mathematics and its applications 27.* Cambridge University Press. MR1015093
[15] Blackwell, D. (1945). The existence of anormal chains. *Bulletin of the American Mathematical Society* **51** 465 – 468. https://doi.org/bams/1183506994. MR0011916
[16] Cavaliere, G. and Georgiev, I. (2020). Inference Under Random Limit Bootstrap Measures. *Econometrica* **88** 2547-2574. MR4192549
[17] Cavaliere, G., Politis, D. N. and Rahbek, A. (2015). Recent Developments in Bootstrap Methods for Dependent Data. *Journal of Time Series Analysis 2015-mar 16 vol. 36 iss. 3* **36**. https://doi.org/10.1111/jtsa.12128. MR3343001
[18] Chen, X. (1999). How Often Does a Harris Recurrent Markov Chain Recur? *The Annals of Probability* **27**. https://doi.org/10.1214/aop/1022677449. MR1733150
[19] Chen, X. (2000). On the limit laws of the second order for additive functionals of Harris recurrent Markov chains. *Probability Theory and Related Fields* **116**. https://doi.org/10.1007/pl00008724. MR1736591
[20] De Coninck, J., Dunlop, F. and Huillet, T. (2009). Random walk versus random line. *Physica A: Statistical Mechanics and its Applications* **388** 4034-4040. https://doi.org/10.1016/j.physa.2009.06.030. MR2552859
[21] Csörgő, S. and Rosalsky, A. (2003). A survey of limit laws for bootstrapped sums. *International Journal of Mathematics and Mathematical Sciences* **2003** 2835–2861. MR2005871
[22] Datta, S. (1996). On asymptotic properties of bootstrap for AR(1) processes. *Journal of Statistical Planning and Inference 1996-aug vol. 53 iss. 3* **53**. https://doi.org/10.1016/0378-3758(95)00147-6. MR1407648
[23] Datta, S. and McCormick, W. P. (1993). Regeneration-Based Bootstrap for Markov Chains. *Canadian Journal of Statistics* **21**. https://doi.org/10.2307/3315810. MR1234760
[24] Delattre, S. and Hoffmann, M. (2002). Asymptotic equivalence for a null recurrent diffusion. *Bernoulli* **8** 139 – 174. MR1895888
[25] Doeblin, W. (1940). Éléments d'une théorie générale des chaînes simples constantes de Markoff. *Annales scientifiques de l'École Normale Supérieure* **57** 61-111. MR0004409
[26] Douc, R., Moulines, E., Priouret, P. and Soulier, P. (2018). *Markov chains. Springer Series in Operations Research and Financial Engineering.* Springer. MR3889011
[27] Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* **7** 1 – 26. https://doi.org/10.1214/aos/1176344552. MR0515681
[28] Franke, J., Kreiss, J.-P. and Mammen, E. (2002). Bootstrap of kernel smoothing in nonlinear time series. *Bernoulli* **8** 1 – 37. MR1884156
[29] Gao, J., Tjøstheim, D. and Yin, J. (2013). Estimation in threshold autoregressive models with a stationary and a unit root regime. *Journal of*

Econometrics **172** 1–13. MR2997126

[30] Gut, A. (2013). *Probability: a graduate course*, 2nd ed ed. *Springer texts in statistics*. Springer. MR2977961

[31] Hall, P. (1990). On the relative performance of bootstrap and Edgeworth approximations of a distribution function. *Journal of Multivariate Analysis* **35** 108-129. https://doi.org/10.1016/0047-259X(90)90019-E. MR1084945

[32] Horowitz, J. L. (2003). Bootstrap Methods for Markov Processes. *Econometrica* **71** 1049-1082. https://doi.org/10.1111/1468-0262.00439. MR1995823

[33] Horowitz, J. L. (2019). Bootstrap Methods in Econometrics. *Annual Review of Economics* **11** 193-224. https://doi.org/10.1146/annurev-economics-080218-025651.

[34] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2 ed. *Grundlehren der mathematischen Wissenschaften 288*. Springer-Verlag Berlin Heidelberg. MR1943877

[35] Jain, N. and Jamison, B. (1967). Contributions to Doeblin's theory of Markov processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **8** 19–40. https://doi.org/10.1007/BF00533942. MR0221591

[36] Kallianpur, H. G. ; Robbins (1954). The sequence of sums of independent random variables. *Duke Mathematical Journal* **21**. https://doi.org/10.1215/S0012-7094-54-02128-6. MR0062976

[37] Karlsen, H. A. and Tjostheim, D. (2001). Nonparametric estimation in null recurrent time series. *The Annals of Statistics* **29**. https://doi.org/10.1214/aos/1009210546. MR1863963

[38] Kasahara, Y. (1984). Limit theorems for Lévy processes and Poisson point processes and their applications to Brownian excursions. *J. Math. Kyoto Univ.* **24** 521–538. https://doi.org/10.1215/kjm/1250521278. MR0766640

[39] Knight, K. (1989). On the Bootstrap of the Sample Mean in the Infinite Variance Case. *The Annals of Statistics* **17** 1168 – 1175. https://doi.org/10.1214/aos/1176347262. MR1015144

[40] Kreiss, J.-P. and Lahiri, S. N. (2012). 1 – Bootstrap Methods for Time Series. In *Time Series Analysis: Methods and Applications. Handbook of Statistics* **30** 3-26. Elsevier. https://doi.org/10.1016/B978-0-444-53858-1.00001-6.

[41] Kreiss, J.-P. and Paparoditis, E. (2003). Autoregressive-Aided Periodogram Bootstrap for Time Series. *The Annals of Statistics* **31** 1923–1955. MR2036395

[42] Kreiss, J.-P. and Paparoditis, E. (2011). Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society* **40** 357-378. https://doi.org/10.1016/j.jkss.2011.08.009. MR2906623

[43] Kreiss, J.-P., Paparoditis, E. and Politis, D. N. (2011). On the range of validity of the autoregressive sieve bootstrap. *The Annals of Statistics* **39** 2103 – 2130. https://doi.org/10.1214/11-AOS900. MR2893863

[44] Kulperger, R. J. and Rao, B. L. S. P. (1989). Bootstrapping a Finite State Markov Chain. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* **51** 178–191. MR1065568
[45] Kunsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics* **17** 1217 – 1241. https://doi.org/10.1214/aos/1176347265. MR1015147
[46] Lahiri, S. (2003). *Resampling methods for dependent data*. Springer Science & Business Media. MR2001447
[47] LePage, R., Woodroofe, M. and Zinn, J. (1981). Convergence to a Stable Distribution Via Order Statistics. *The Annals of Probability* **9** 624 – 632. https://doi.org/10.1214/aop/1176994367. MR0624688
[48] MacKinnon, J. G. (2006). Bootstrap Methods in Econometrics*. *Economic Record* **82** S2-S18. https://doi.org/10.1111/j.1475-4932.2006.00328.x.
[49] Malinovskii, V. K. (1987). Limit Theorems for Harris Markov Chains, I. *Theory of Probability and Its Applications* **31**. https://doi.org/10.1137/1131033. MR0850991
[50] Meyn, S., Tweedie, R. and Glynn, P. (2009). *Markov chains and stochastic stability*, 2 ed. *Cambridge Mathematical Library*. Cambridge University Press. MR2509253
[51] Myklebust, T., Karlsen, H. A. and Tjøstheim, D. (2012). Null Recurrent Unit Root Process. *Econometric Theory* **28**. https://doi.org/10.1017/S0266466611000119. MR2899213
[52] Nummelin, E. (1978). A splitting technique for Harris recurrent Markov chains. *Probability Theory and Related Fields* **43**. https://doi.org/10.1007/bf00534764. MR0501353
[53] Nummelin, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators. Cambridge Tracts in Mathematics 83*. Cambridge University Press. MR0776608
[54] Palm, F. C., Smeekes, S. and Urbain, J.-P. (2008). Bootstrap unit-root tests: comparison and extensions. *Journal of Time Series Analysis* **29** 371–401. MR2392778
[55] Paparoditis, E. (2002). *Frequency Domain Bootstrap for Time Series* In *Empirical Process Techniques for Dependent Data* 365–381. Birkhäuser Boston, Boston, MA. https://doi.org/10.1007/978-1-4612-0099-4_14. MR1958790
[56] Paparoditis, E. and Politis, D. N. (2000). *The Continuous-Path Block-Bootstrap* In *Papers in Honor of George Gregory Roussas* 305–320. De Gruyter, Berlin, Boston. https://doi.org/doi:10.1515/9783110942002-021.
[57] Paparoditis, E. and Politis, D. N. (2001). Tapered Block Bootstrap. *Biometrika* **88** 1105–1119. MR1872222
[58] Paparoditis, E. and Politis, D. N. (2001). A Markovian Local Resampling Scheme for Nonparametric Estimators in Time Series Analysis. *Econometric Theory* **17** 540–566. MR1841820
[59] Paparoditis, E. and Politis, D. N. (2002). Local block bootstrap.

Comptes Rendus Mathematique **335** 959-962. https://doi.org/10.1016/S1631-073X(02)02578-5. MR1952557

[60] Paparoditis, E. and Politis, D. N. (2002). The local bootstrap for Markov processes. *Journal of Statistical Planning and Inference* **108** 301-328. C.R. Rao 80th Birthday Felicitation Volume, Part II. https://doi.org/10.1016/S0378-3758(02)00315-4. MR1947405

[61] Politis, D. N. and Romano, J. P. (1991). *A circular block-resampling procedure for stationary data*. Purdue University. Department of Statistics. MR1197789

[62] Politis, D. N. and Romano, J. P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association* **89** 1303–1313. MR1310224

[63] Pollard, D. (1984). *Convergence of stochastic processes*. Springer. MR0762984

[64] Rajarshi, M. B. (1990). Bootstrap in Markov-sequences based on estimates of transition density. *Annals of the Institute of Statistical Mathematics* **42** 253–268. MR1064787

[65] Revesz, P. (2005). *Random Walk in Random and Non-Random Environments*, 2nd ed ed. World Scientific. MR2168855

[66] Tang, X.-S., Huang, H.-B., Liu, X.-F., Li, D.-Q. and Liu, Y. (2023). Efficient Bayesian method for characterizing multiple soil parameters using parametric bootstrap. *Computers and Geotechnics* **156** 105296. https://doi.org/10.1016/j.compgeo.2023.105296.

[67] Zoubir, A. M. and Boashash, B. (1998). The bootstrap and its application in signal processing. *IEEE Signal Processing Magazine* **15** 56-76. https://doi.org/10.1109/79.647043.

[68] Zoubir, A. M. and Iskander, D. R. (2004). *Bootstrap Techniques for Signal Processing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511536717.