Markov Processes Relat. Fields **9**, 413–468 (2003)

Extended Renovation Theory and Limit Theorems for Stochastic Ordered Graphs*

S. Foss[1] **and T. Konstantopoulos**[2]

[1] Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh EH14 4AS, UK. E-mail: S.Foss@ma.hw.ac.uk
[2] Dept. of Electr. & Computer Eng., University of Texas at Austin, Austin, TX 78712, USA and Dept. of Mathematics, University of Patras, 26500 Patras, Greece.
E-mail: tk@math.upatras.gr

Received July 11, 2002, revised February 17, 2003

Abstract. We extend Borovkov's renovation theory to obtain criteria for coupling-convergence of stochastic processes that do not necessarily obey stochastic recursions. The results are applied to an "infinite bin model", a particular system that is an abstraction of a stochastic ordered graph, i.e., a graph on the integers that has (i, j), $i < j$, as an edge, with probability p, independently from edge to edge. A question of interest is an estimate of the length L_n of a longest path between two vertices at distance n. We give sharp bounds on $C = \lim_{n\to\infty}(L_n/n)$. This is done by first constructing the unique stationary version of the infinite bin model, using extended renovation theory. We also prove a functional law of large numbers and a functional central limit theorem for the infinite bin model. Finally, we discuss perfect simulation, in connection to extended renovation theory, and as a means for simulating the particular stochastic models considered in this paper.

Keywords: stationary and ergodic processes, renovation theory, functional limit theorems, weak convergence, coupling, perfect simulation

AMS Subject Classification: Primary 60G10, 60F17; Secondary 68U20, 05C80

*This work was supported in part by NSF grant ANI-9903495, the INTAS-00-265 project, and a Carathéodory research award

Contents

1. Introduction

A large fraction of the applied probability literature is devoted to the study of existence and uniqueness of a stationary solution to a stochastic dynamical system, as well as convergence toward such a stationary solution. Examples abound in several application areas, such as queueing theory, stochastic control, simulation algorithms, etc. Often, the model studied possesses a Markovian property, in which case several classical tools are available. In the absence of Markovian property, one has few tools to rely on, in general. The concept of *renovating event* was introduced by Borovkov [6] in an attempt to produce general conditions for a strong type of convergence of a stochastic process, satisfying a stochastic recursion, to a stationary process. Other general conditions for existence/uniqueness questions can be found, e.g., in [1, 2] and [15]. The so-called *renovation theory* or *method of renovating events* has a flavor different from the aforementioned papers, in that it leaves quite a bit of freedom in the choice of a renovating event, which is what makes it often hard to apply: some ingenuity is required in constructing renovating events. Nevertheless, renovation theory has found several applications, especially in queueing-type problems (see, e.g., [11] and [3] for a variety of models and techniques in this area).

Renovation theory is stated for "stochastic recursive processes", i.e., random sequences $\{X_n\}$ defined by recursive relations of the form

$$X_{n+1} = f(X_n, \xi_{n+1}),$$

where f is appropriately measurable function, and $\{\xi_n\}$ a stationary random sequence. In this paper, we take a fresh look at renovation theory and formulate it for processes that do not necessarily obey stochastic recursions. We use the terminology *extended renovation theory*. Our objective here is threefold: first, we give a self-consistent overview of (extended) renovation theory, and strong coupling notions; second, we present simple proofs of renovating criteria (Theorem 3.1); third, we shed some light into the so-called *coupling from the past* property, which has drawn quite a bit of attention recently, especially in connection to the Propp – Wilson algorithm (see [22]) for perfect simulation (alternative terminology: exact sampling). We do this, by defining strong forward and backward coupling times. We also pay particular attention to a special type of backward coupling times, those that we call *verifiable times*: they form precisely the class of those times that can be simulated. Verifiable backward times exist, e.g., for irreducible Markov chains with finite state space (and hence exact sampling from the stationary distribution is possible), but they also exist in other models, such as the ones presented in the second part of the paper.

A useful model is what we call *infinite bin model*. It is presented as an abstraction of a *stochastic ordered graph*. The latter model appears in mathematical ecology (it models community food webs; see, e.g., [21] and [13]), and in performance evaluation of computer systems (it models task graphs; see, e.g., [18]). The model, in its simplest form, is a graph on the integers $\mathbf{Z}$, where edges (always directed to the right) appear independently at random with probability p. A question of interest is to estimate the length L_n of the longest path between two vertices at distance n. Other questions are to estimate the number of paths with longest length. To answer these questions, we turn the graph into an infinite bin model, by considering the order statistics of the paths of various lengths. The infinite bin model (which is a system of interest in its own) consists of infinite number of bins arranged on the line and indexed, by the non-positive integers $\mathbf{Z}_- := \{0, -1, -2, \ldots\}$. The bin in position $-k \in \mathbf{Z}_-$ contains, at time n, a finite number of particles, denoted by $X_n(-k)$. The state of the system is $X_n = [\ldots, X_n(-k), \ldots, X_n(0)]$, an element of $\mathbf{N}^{\mathrm{Z}_-}$. We refer to the set $\mathbf{N}^{\mathrm{Z}_-}$ as the state space, or the space of configurations. To create X_{n+1}, precisely one particle of the current configuration X_n is chosen in some random manner. If the particle is in bin $-k \leq 1$, then a new particle is created and placed in bin $-k+1$. Otherwise, if the chosen particle is in bin 0 then a new bin is created to hold the child particle and a relabeling of the bins occurs: the existing ones are shifted by one place to the left (and are re-indexed) and the new bin is given the label 0. Despite the fact that X_n is a stochastic recursive sequence, we cannot apply the usual renovation theory directly to obtain a stationary version.

Rather, the extended version of the theory (developed in the first part of the paper) should be applied.

We thus achieve the following, by applying extended renovation theory to the infinite bin model: each finite-dimensional projection of the state of the system strongly couples with a stationary version. However, the state itself does not. We thus obtain Theorem 7.1 which states that the infinite bin model has a unique stationary solution $\{\tilde{X}_n\}$, and, as $n \to \infty$, the law of X_n converges to the law of $\tilde{X}_0$, *weakly* (but without strong coupling). Whereas the meaning of X_n is clear in terms of the stochastic ordered graph, the meaning of $\tilde{X}_n$ is not. It is a useful abstract stochastic process which we use in order to obtain meaningful estimates for the longest length sequence L_n: there is a deterministic constant C, such that, as $n \to \infty$,

$$L_n/n \to C, \quad \text{a.s.},$$

where C depends on the connectivity probability p (or the non-connectivity probability $q = 1 - p$). In Theorem 10.1, we give explicit upper and lower bounds on $C(q)$ for all q, and, as a corollary, we obtain good asymptotics for $C(q)$ when q is small, and when q is large. More precisely, we find that

$$C = \begin{cases} 1 - q + q^2 - 3q^3 + 7q^4 + O(q^5), & \text{as } q \to 0 \quad \text{(heavy graph)}, \\ O((1-q)\log(1-q)), & \text{as } q \to 1 \quad \text{(sparse graph)}. \end{cases}$$

In addition, we present a functional law of large numbers for the infinite bin model (Theorem 8.1) which states that

$$\frac{1}{n} \sum_{k=0}^{[nCt]} X_n(-k) \to t, \qquad \text{as } n \to \infty\ , \qquad \text{uniformly in } t \in [0,1], \text{ a.s.}$$

We complement this result by a corresponding central limit theorem (Theorem 9.1), stating that

$$\left\{ \sqrt{n} \left(\frac{1}{n} \sum_{k=0}^{[nCt]} X_n(-k) - t \right) \right\}_{t \geq 0} \Rightarrow \text{Brownian Motion}, \qquad \text{as } n \to \infty\ ,$$

where $\Rightarrow$ denotes weak convergence in $D[0, \infty)$, with the topology of uniform convergence on compacta, provided that we introduce some independence assumptions on the stochastic recursion for the infinite bin model. If the scaling is slightly changed, namely if, instead of $[nCt]$ at the upper limit of the summation, we consider $[L_n t]$ (a change which, in a sense, is small due to the fact that $L_n = nC + o(n)$, a.s.), we obtain a different result

$$\left\{ \sqrt{n} \left(\frac{1}{n} \sum_{k=0}^{[L_n t]} X_n(-k) - t \right) \right\}_{0 \leq t \leq 1} \Rightarrow \text{Brownian Bridge}, \qquad \text{as } n \to \infty\ .$$

The two functional central limit theorems hold both for the stationary version of the infinite bin model as well as the infinite bin model that starts from a trivial initial state (transient infinite bin model). From a physical point of view, the latter scaling is more natural for the transient infinite bin model.

The paper is organized as follows: Section 2 is an overview of the coupling notions and contains, in particular, a proof of the equivalence between forward and backward coupling (Theorem 2.1). Section 3 gives a short account of the extended renovation theory and its main criterion (Theorem 3.1). Section 4 defines verifiable times and gives a criterion for verifiability (Theorem 4.1). In addition, we present a so-called perfect simulation algorithm that works in a rather general setup, provided that renovation events of special type exist. The next two Sections, 5 and 6, deal with specializing the extended renovation theory to stochastic recursive sequences and to functionals of them (thus justifying the terminology "extended"). The infinite bin model is introduced in Section 7. Its stationarity and convergence properties are dealt with in the same section, by an application of the extended renovation theory. Sections 8 and 9 describe the functional law of large numbers and central limit theorem, respectively, for the infinite bin model. Section 10 presents the application of the above to stochastic ordered graphs, describes their relation to infinite bin models, and gives bounds and asymptotics on the constant C. Finally, Section 11 presents an application in queueing theory and open problems.

Before closing this section, a word of caution is due. In this paper, we differentiate a *stationary version* from a *stationary solution*. When a stochastic process X couples with a stationary process $\tilde{X}$ (see Definition 2.1), we refer to $\tilde{X}$ as the *stationary version* of X. Immediately, we have that there can be only one such $\tilde{X}$. On the other hand, when we deal with a stochastic recursion (with stationary driver), we talk about a *stationary solution* of the stochastic recursion whenever we have a process $\tilde{X}$ that is simultaneously stationary *and* satisfies the stochastic recursion. Such stationary solutions may be many, which may or may not be stationary versions of particular solutions of the stochastic recursion.

2. Strong coupling notions

To prepare for the extended renovation theory, we start by defining the notions of coupling (and coupling convergence) that we need. For general notions of coupling we refer to the monographs of Lindvall [19] and Thorisson [23]. For the strong coupling notions of this paper, we refer to [9].

Consider a sequence of random variables $\{X_n, n \in \mathbf{Z}\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and taking values in another measurable space $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$. We study various ways according to which X couples with another stationary process $\{\tilde{X}_n, n \in \mathbf{Z}\}$. We push the stationarity structure into the probability space itself, by assuming the existence of a flow (i.e., a measurable bijection)

$\theta : \Omega \to \Omega$ that leaves the probability measure P invariant, i.e., $\mathsf{P}(\theta^k A) = \mathsf{P}(A)$, for all $A \in \mathcal{F}$, and all $k \in \mathbf{Z}$. In this setup, a stationary process $\{\tilde{X}_n\}$ is, by definition, a θ-compatible process in the sense that $\tilde{X}_{n+1} = \tilde{X}_n \circ \theta$ for all $n \in \mathbf{Z}$. Likewise, a sequence of events $\{A_n\}$ is stationary iff their indicator functions $\{\mathbf{1}_{A_n}\}$ are stationary. Note that, in this case, $\mathbf{1}_{A_n} \circ \theta = \mathbf{1}_{\theta^{-1} A_n} = \mathbf{1}_{A_{n+1}}$ for all $n \in \mathbf{Z}$. In order to avoid technicalities, we assume that the σ-algebra $\mathcal{B}_{\mathcal{X}}$ is countably generated. The same assumption, without special notice, will be made for all σ-algebras below.

We next present three notions of coupling: *simple coupling*, *strong* (*forward*) *coupling* and *backward coupling*. To each of these three notions there corresponds a type of convergence. These are called *c-convergence*, *sc-convergence*, and *bc-convergence*, respectively. The definitions below are somewhat formal by choice: there is often a danger of confusion between these notions. To guide the reader, we first present an informal discussion. Simple coupling between two processes (one of which is usually stationary) refers to the fact that the two processes are a.s. identical, eventually. To define strong (forward) coupling, consider the family of processes that are derived from X "started from all possible initial states at time 0". To explain what the phrase in quotes means in a non-Markovian setup, place the origin of time at the negative index $-m$, and run the process forward till a random state at time 0 is reached: this is the process X^{-m} formally defined in (2.2). Strong coupling requires the existence of a finite random time $\sigma \geq 0$ such that all these processes are identical after σ. Backward coupling is — in a sense — the dual of strong coupling: instead of fixing the starting time (time 0) and waiting till the random time σ, we play a similar game with a random starting time (time $-\tau \leq 0$) and wait till coupling takes place at a fixed time (time 0). That is, backward coupling takes place if there is a finite random time $-\tau \leq 0$ such that all the processes started at times prior to $-\tau$ are coupled forever after time 0. The main theorem of this section (Theorem 2.1) says that strong (forward) coupling and backward coupling are equivalent, whereas an example (Example 2.1) shows that they are both strictly stronger than simple coupling.

We first consider simple coupling. Note that our definitions are more general than usual because we do not necessarily assume that the processes are solutions of stochastic recursions.

Definition 2.1 (simple coupling).

1) The minimal coupling time between X and $\tilde{X}$ is defined by

$$\nu = \inf\{n \geq 0 : \ \forall\, k \geq n \ \ X_k = \tilde{X}_k\}.$$

2) More generally, a random variable ν' is said to be a coupling time between X and $\tilde{X}$ iff[1]

$$X_n = \tilde{X}_n, \qquad \text{a.s. on } \{n \geq \nu'\}.$$

[1] "B a.s. on A" means $\mathsf{P}(A - B) = 0$.

3) We say that X coupling-converges (or c-converges) to $\tilde{X}$ iff $\nu < \infty$, a.s., or, equivalently, if $\nu' < \infty$, a.s., for some coupling time ν'.

Notice that the reason we call ν "minimal" is because (i) it is a coupling time, and (ii) any random variable ν' such that $\nu' \geq \nu$, a.s., is also a coupling time.

Proposition 2.1 (c-convergence criterion). *X c-converges to $\tilde{X}$ iff*

$$\mathsf{P}(\liminf_{n\to\infty}\{X_n = \tilde{X}_n\}) = 1.$$

Proof. It follows from the equality

$$\{\nu < \infty\} = \bigcup_{n\geq 0} \bigcap_{k\geq n} \{X_k = \tilde{X}_k\},$$

the right-hand side of which is the event $\liminf_{n\to\infty}\{X_n = \tilde{X}_n\}$. □

It is clear that c-convergence implies convergence in total variation, i.e.,

$$\lim_{n\to\infty} \sup_{B\in\mathcal{B}_{\mathcal{X}}^\infty} \Big|\mathsf{P}\big((X_n, X_{n+1}, \ldots) \in B\big) - \mathsf{P}\big((\tilde{X}_n, \tilde{X}_{n+1}, \ldots) \in B\big)\Big| = 0, \qquad (2.1)$$

simply because the left-hand side is dominated by $\mathsf{P}(\nu \geq n)$ for all n. In fact, the converse is also true, viz., (2.1) implies c-convergence (see [23], Theorem 9.4). Thus, c-convergence is a very strong notion of convergence, but not the strongest one that we are going to deal with in this paper.

The process $\tilde{X}$ in (2.1) will be referred to as the *stationary version* of X. Note that the terminology is slightly non-standard because, directly from the definition, if such a $\tilde{X}$ exists, it is automatically unique (due to coupling). The term is usually defined for stochastic recursive sequences (SRS). To avoid confusion, we talk about a *stationary solution* of an SRS, which may not be unique. See Section 5 for further discussion.

A comprehensive treatment of the notions of coupling, as well as the basic theorems and applications can be found in [9], for the special case of processes which form stochastic recursive sequences. For the purposes of our paper, we need to formulate some of these results beyond the SRS realm, and this is done below.

It is implicitly assumed above (see the definition of ν) that 0 is the "origin of time". This is, of course, totally arbitrary. We now introduce the notation

$$X_n^{-m} := X_{m+n} \circ \theta^{-m}, \quad m \geq 0,\ n \geq -m,$$

and consider the family of processes

$$X^{-m} := (X_0^{-m}, X_1^{-m}, \ldots), \quad m = 0, 1, \ldots, \qquad (2.2)$$

and the minimal coupling time $\sigma(m)$ of X^{-m} with $\tilde{X}$. The definition becomes clearer when X itself is a SRS (see Section 5).

Definition 2.2 (strong coupling).
1) The minimal strong coupling time between X and $\tilde{X}$ is defined by

$$\sigma = \sup_{m \geq 0} \sigma(m), \quad \text{where}$$
$$\sigma(m) = \inf\{n \geq 0 : \ \forall\, k \geq n, \quad X_k^{-m} = \tilde{X}_k\}.$$

2) More generally, a random variable σ' is said to be a strong coupling time (or sc-time) between X and $\tilde{X}$ iff

$$\tilde{X}_n = X_n^0 = X_n^{-1} = X_n^{-2} = \cdots, \qquad \text{a.s. on } \{n \geq \sigma'\}.$$

3) We say that $\{X_n\}$ strong-coupling-converges (or sc-converges) to $\{\tilde{X}_n\}$ iff $\sigma < \infty$, a.s.

Again, it is clear that the minimal strong coupling time σ is a strong coupling time, and that any σ' such that $\sigma' \geq \sigma$, a.s., is also a strong coupling time.

Even though strong coupling is formulated by means of two processes, X and a stationary $\tilde{X}$, we will see that the latter is not needed in the definition.

Example 2.1 (see [9]). We now give an example to show the difference between coupling and strong coupling. Let $\{\xi_n, n \in \mathbf{Z}\}$ be an i.i.d. sequence of random variables with values in $\mathbf{Z}_+$ such that $\mathsf{E}\xi_0 = \infty$. Let

$$X_n = (\xi_0 - n)^+, \quad \tilde{X}_n = 0, \quad n \in \mathbf{Z}.$$

The minimal coupling time between $(X_n, n \geq 0)$ and $(\tilde{X}_n, n \geq 0)$ is $\nu = \xi_0 < \infty$, a.s. Hence $\tilde{X}$ is the stationary version of X. Since

$$X_n^{-m} = X_{m+n} \circ \theta^{-m} = (\xi_{-m} - (m+n))^+,$$

the minimal coupling time between $(X_n^{-m}, n \geq 0)$ and $(\tilde{X}_n, n \geq 0)$ is $\sigma(m) = (\xi_{-m} - m)^+$. Hence the minimal *strong* coupling time between X and $\tilde{X}$ is $\sigma = \sup_{m \geq 0} \sigma(m)$. But $\mathsf{P}(\sigma \leq n) = \mathsf{P}(\forall m \geq 0, \quad \xi_m - m \leq n) = \prod_{m \geq 0} \mathsf{P}(\xi_0 \leq m+n)$, and, since $\sum_{j \geq 0} \mathsf{P}(\xi_0 > j) = \infty$, we have that the latter infinite product is zero, i.e., $\sigma = +\infty$, a.s. So, even though X couples with $\tilde{X}$, it does not couple strongly.

Proposition 2.2 (sc-convergence criterion). *X sc-converges to $\tilde{X}$ iff*

$$\mathsf{P}\Big(\liminf_{n \to \infty} \bigcap_{m \geq 0} \{\tilde{X}_n = X_n^{-m}\}\Big) = 1.$$

Proof. It follows from the definition of σ that

$$
\begin{aligned}
\{\sigma < \infty\} = \bigcup_{n\geq 0}\{\sigma \leq n\} &= \bigcup_{n\geq 0}\bigcap_{m\geq 0}\{\sigma(m)\leq n\} \\
&= \bigcup_{n\geq 0}\bigcap_{m\geq 0}\bigcap_{k\geq n}\{\tilde{X}_k = X_k^{-m}\} \\
&= \bigcup_{n\geq 0}\bigcap_{k\geq n}\bigcap_{m\geq 0}\{\tilde{X}_k = X_k^{-m}\} \\
&= \liminf_{n\to\infty}\bigcap_{m\geq 0}\{\tilde{X}_n = X_n^{-m}\},
\end{aligned}
$$

and this proves the claim. □

The so-called backward coupling (see [9,16] for this notion in the case of SRS) is introduced next. This does not require the stationary process $\tilde{X}$ for its definition. Rather, the stationary process is constructed once backward coupling takes place. Even though the notion appears to be quite strong, it is not infrequent in applications.

Definition 2.3 (backward coupling).
1) The minimal backward coupling time for the random sequence $\{X_n, n \in \mathbf{Z}\}$ is defined by $\tau = \sup_{m\geq 0}\tau(m)$, where

$$\tau(m) = \inf\big\{n\geq 0:\ \forall k\geq 0,\quad X_m^{-n} = X_m^{-(n+k)}\big\}.$$

2) More generally, we say that τ' is a backward coupling time (or bc-time) for X iff

$$\forall m\geq 0,\quad X_m^{-t} = X_m^{-(t+1)} = X_m^{-(t+2)} = \cdots, \qquad \text{a.s. on } \{t\geq \tau'\}.$$

3) We say that $\{X_n\}$ backward-coupling converges (or bc-converges) iff $\tau < \infty$, a.s.

Note that τ is a backward coupling time and that any τ' such that $\tau' \geq \tau$, a.s., is a backward coupling time. We next present the equivalence theorem between backward and forward coupling.

Theorem 2.1 (coupling equivalence). *Let τ be the minimal backward coupling time for X. There is a stationary process $\tilde{X}$ such that the strong coupling time σ between X and $\tilde{X}$ has the same distribution as τ on $\mathbf{Z}_+\cup\{+\infty\}$. Furthermore, if $\tau < \infty$ a.s., then $\tilde{X}$ is the stationary version of X.*

Proof. Using the definition of τ, we write

$$\begin{aligned}\{\tau < \infty\} &= \bigcup_{n\geq 0}\{\tau \leq n\} = \bigcup_{n\geq 0}\bigcap_{m\geq 0}\{\tau(m) \leq n\} \\ &= \bigcup_{n\geq 0}\bigcap_{m\geq 0}\bigcap_{\ell\geq n}\{X_m^{-n} = X_m^{-\ell}\} = \bigcup_{n\geq 0}\bigcap_{\ell\geq n}\bigcap_{m\geq 0}\{X_m^{-n} = X_m^{-\ell}\}. \qquad (2.3)\end{aligned}$$

Consider, as in (2.2), the process $X^{-n} = (X_0^{-n}, X_1^{-n}, X_2^{-n}, \ldots)$, with values in $\mathcal{X}^{\mathrm{Z}_+}$. Using this notation, (2.3) can be written as

$$\begin{aligned}\{\tau < \infty\} &= \bigcup_{n\geq 0}\bigcap_{\ell\geq n}\{X^{-n} = X^{-\ell}\} \\ &= \{\exists n \geq 0,\ X^{-n} = X^{-(n+1)} = X^{-(n+2)} = \cdots\}.\end{aligned}$$

Thus, on the event $\{\tau < \infty\}$, the random sequence X^{-n} is equal to some fixed random element of $\mathcal{X}^{\mathrm{Z}_+}$ for all large n (it is eventually a constant sequence). Let $\tilde{X} = (\tilde{X}_0, \tilde{X}_1, \ldots)$ be this random element; it is defined on $\{\tau < \infty\}$. Let ∂ be an arbitrary fixed member of $\mathcal{X}^{\mathrm{Z}_+}$ and define $\tilde{X} \equiv \partial$ outside $\{\tau < \infty\}$. Since the event $\{\tau < \infty\}$ is a.s. invariant under θ^n, for all $n \in \mathbf{Z}$, we obtain that $\tilde{X}$ is a stationary process. Let σ be the strong coupling time between X and $\tilde{X}$. It is easy to see that, for all $n \geq 0$,

$$\{\sigma\circ\theta^{-n} \leq n\} = \bigcap_{\ell\geq n}\{\tilde{X} = X^{-\ell}\} = \{\tau \leq n\}. \qquad (2.4)$$

Indeed, on one hand, from the definition of τ, we have

$$\{\tau \leq n\} = \bigcap_{\ell\geq n}\{X^{-n} = X^{-\ell}\}.$$

Now, using the $\tilde{X}$ we just defined we can write this as

$$\{\tau \leq n\} = \bigcap_{\ell\geq n}\{\tilde{X} = X^{-\ell}\}. \qquad (2.5)$$

On the other hand, from the definition of σ (that is, the strong coupling time between X and $\tilde{X}$), we have

$$\{\sigma \leq n\} = \bigcap_{k\geq 0}\{\tilde{X}_{n+k} = X_{n+k}^0 = X_{n+k}^{-1} = X_{n+k}^{-2} = \cdots\}.$$

Applying a shifting operation on both sides,

$$\{\sigma\circ\theta^{-n} \leq n\} = \bigcap_{k\geq 0}\{\tilde{X}_k = X_k^n = X_k^{n-1} = X_k^{n-2} = \cdots\}. \qquad (2.6)$$

The events on the right-hand sides of (2.5) and (2.6) are identical. Hence (2.4) holds for all n, and thus $\mathsf{P}(\tau \le n) = \mathsf{P}(\sigma \circ \theta^n \le n) = \mathsf{P}(\sigma \le n)$, for all n. Finally, if $\mathsf{P}(\tau < \infty) = 1$ then $\mathsf{P}(\sigma < \infty) = 1$, and this means that X sc-converges to $\tilde{X}$. In particular, we have convergence in total variation, and so $\tilde{X}$ is the stationary version of X. □

Corollary 2.1. *The following statements are equivalent:*

1. X bc-converges;
2. $\lim_{n\to\infty} \mathsf{P}\big(\forall m \ge 0, \quad X_m^{-n} = X_m^{-(n+1)} = X_m^{-(n+2)} = \cdots\big) = 1$;
3. X sc-converges;
4. $\lim_{n\to\infty} \mathsf{P}\big(\forall k \ge 0,\ X_{n+k}^{0} = X_{n+k}^{-1} = X_{n+k}^{-2} = \cdots\big) = 1$.

We can view any of the equivalent statements of Corollary 2.1 as an "intrinsic criterion" for the existence the stationary version of X.

Corollary 2.2. *Suppose that X bc-converges and let τ be the minimal backward coupling time. Let $\tilde{X}_0 = X_\tau \circ \theta^{-\tau}$. Then $\tilde{X}_n = \tilde{X}_0 \circ \theta^n$ is the stationary version of X. Furthermore, if τ' is any a.s. finite backward coupling time then $X_\tau \circ \theta^{-\tau} = X'_\tau \circ \theta^{-\tau'}$, a.s.*

Proof. Let $\tilde{X}$ be the stationary version of X. It follows, from the construction of $\tilde{X}$ in the proof of Theorem 2.1, that

$$(X_0^{-t}, X_1^{-t}, X_2^{-t}, \ldots) = (\tilde{X}_0, \tilde{X}_1, \tilde{X}_2, \ldots), \qquad \text{a.s. on } \{t \ge \tau\}. \tag{2.7}$$

Thus, in particular, $\tilde{X}_0 = X_0^{-t} = X_t \circ \theta^{-t}$, a.s. on $\{t \ge \tau\}$. Since $\mathsf{P}(\tau < \infty) = 1$, it follows that $\tilde{X}_0 = X_\tau \circ \theta^{-\tau}$, a.s. Now, if τ' is any backward coupling time, then (2.7) is true with τ' in place of τ; and if $\tau' < \infty$, a.s., then, as above, we conclude that $\tilde{X}_0 = X_{\tau'} \circ \theta^{-\tau'}$. □

3. Extended renovation theory

In this section we extend the theory of renovating events in a rather general setup. The method of renovating events for stochastic recursive sequences was introduced by Borovkov [6, 7] and further developed by Foss [16] and Borovkov and Foss [9, 10]; see also the monographs by Borovkov [7] and [8, Chapter 3]. We present a generalization of this theory, that is applicable to stochastic processes that are not necessarily stochastic recursive sequences. In addition to its generality, the setup below can be formulated quite simply and has applications to the models we are considering.

As before, consider an arbitrary process $\{X_n, n \in \mathbf{Z}\}$ with values in $(\mathcal{X}, \mathcal{B}_\mathcal{X})$. We seek a sufficient criterion for its bc-convergence. A doubly-indexed stationary sequence $\{H_{m,n}, \ -\infty < m \leq n < \infty\}$, with values in the same space, satisfies, by definition,

$$H_{m,n} \circ \theta^j = H_{m+j,n+j}, \quad \forall j \in \mathbf{Z}.$$

We may call such a process *"stationary background"*. In its most general form, the "renovating events criterion" is

Theorem 3.1 (extended renovation theorem). *Let $\{X_n, n \in \mathbf{Z}\}$ be a sequence of random variables with values in $(\mathcal{X}, \mathcal{B}_\mathcal{X})$, defined on $(\Omega, \mathcal{F}, \mathsf{P}, \theta)$, where θ is a P-preserving ergodic flow. Suppose there exists stationary sequence of events $\{A_n, n \in \mathbf{Z}\}$ with $\mathsf{P}(A_0) > 0$, a stationary background $\{H_{m,n}, \ -\infty < m \leq n < \infty\}$, and an index $n_0 \in \mathbf{Z}$, such that*

$$\forall n \geq n_0, \quad \forall j \geq 0, \quad X_{n+j} = H_{n,n+j}, \quad \text{a.s. on } A_n. \tag{3.1}$$

Then $\{X_n\}$ bc-converges.

N.B. The events A_n for which (3.1) holds are called *renovating events*.

Proof of Theorem 3.1. Without loss of generality assume $n_0 = 0$, and write the condition of the theorem symbolically as

$$\forall n \geq 0, \quad \forall j \geq 0, \qquad X_{n+j} \mathbf{1}_{A_n} = H_{n,n+j} \mathbf{1}_{A_n}, \qquad \text{a.s.}$$

Since the above statement holds almost surely, and since θ preserves P, we also have

$$\forall n, j \geq 0, \quad \forall m \in \mathbf{Z}, \qquad X_{n+j} \circ \theta^{-m} \mathbf{1}_{A_{n-m}} = H_{n-m,n-m+j} \mathbf{1}_{A_{n-m}}, \qquad \text{a.s.} \tag{3.2}$$

Put $i = m-n$, $k = n-m+j$. Observe that this change of indices transforms the set $\{(n,j,m) \in \mathbf{Z}^3 : \ n \geq 0, j \geq 0\}$ into the set $\{(k,i,m) \in \mathbf{Z}^3 : \ m \geq i \geq -k\}$. So (3.2) is written as

$$\forall m \geq i \geq -k, \qquad X_k^{-m} \mathbf{1}_{A_{-i}} = H_{-i,k} \mathbf{1}_{A_{-i}}, \qquad \text{a.s.},$$

where we also used the notation $X_k^{-m} = X_{m+k} \circ \theta^{-m}$. Define now the random time

$$\gamma := \inf\{i \geq 0 : \ \mathbf{1}_{A_{-i}} = 1\}. \tag{3.3}$$

Since $\mathsf{P}(A_0) > 0$ and θ is ergodic, we have $\gamma < \infty$, a.s. Thus,

$$\forall \ m \in \mathbf{Z}, \ \ \forall \ k \geq 0, \qquad H_{-\gamma,k} = X_k^{-m}, \qquad \text{a.s. on } \{m \geq \gamma\}.$$

Since, obviously, $\{m \ge \gamma\} \subseteq \{m+1 \ge \gamma\}$ for all m, we also have the seemingly stronger statement

$$\forall k \ge 0, \quad H_{-\gamma,k} = X_k^{-m}, \quad H_{-\gamma,k} = X_k^{-(m+1)}, \ldots, \quad \text{a.s. on } \{m \ge \gamma\}. \tag{3.4}$$

This implies that, for all $m \ge 0$,

$$\mathsf{P}(m \ge \gamma) \le \mathsf{P}\big(\forall\ k \ge 0,\ H_{-\gamma,k} = X_k^{-m} = X_k^{-(m+1)} = \cdots\big). \tag{3.5}$$

But the left side tends to 1 as $m \to \infty$, and so we can conclude by applying Corollary 2.1. □

Corollary 3.1. *Under the conditions of Theorem* 3.1*, the random variable* γ *defined in* (3.3) *is a* (*not necessarily minimal*) *backward coupling time. Furthermore,* X *has the stationary version* $\tilde{X}$ *given by*

$$\tilde{X}_n = X_{\gamma+n} \circ \theta^{-\gamma} = H_{-\gamma,n}.$$

Proof. That γ is a backward coupling time follows from (3.4). That the stationary version of X is given by $\tilde{X}_n = X_{\gamma+n} \circ \theta^{-\gamma}$ follows from Corollary 2.2. That the stationary version is also given by $\tilde{X}_n = H_{-\gamma,n}$ follows from (3.4) and the way that $\tilde{X}$ was constructed in the proof of Theorem 2.1. □

It is often useful to consider, instead of single random index γ, as in (3.3), for which $\mathbf{1}_{A_{-\gamma}} = 1$, a.s., the entire sequence $\{j \in \mathbf{Z} : \mathbf{1}_{A_j} = 1\}$ and enumerate it as follows

$$\cdots < -\gamma_{-1} < -\gamma_0 \le 0 < \gamma_1 < \gamma_2 < \cdots$$

In this notation, $\gamma_0 = \gamma$. This sequence is nonterminating on both sides and it is clear that the stationary version can be constructed by starting from any term of this sequence. Namely, for any $r \ge 1$, we have that $H_{\gamma_r,n} = \tilde{X}_n$ for all $n \ge \gamma_r$, a.s.

4. The concept of verifiability and perfect simulation

One application of the theory is the simulation of stochastic systems. If we could sample the process at a bc-time, then would actually be simulating its stationary version. This is particularly useful in *Markov Chain Monte Carlo* applications. Recently, Propp and Wilson [22] used the so-called *perfect simulation method* for the simulation of the invariant measure of a Markov chain. The method is actually based on sampling at a bc-time. To do so, however, one must be able to generate a bc-time from a finite history of the process. In general, this may not be possible because, even in the case when suitable renovation events can be found, they may depend on the entire history of the process.

We are thus led to the concept of a *verifiable time*. Its definition, given below, requires introducing a family of σ-fields $\{\mathcal{G}_{-j,m}, -j \leq 0 \leq m\}$, such that $\mathcal{G}_{-j,m}$ increases if j or m increases. We call this simply an increasing family of σ-fields. For fixed m, a *backwards stopping time* $\tau \geq 0$ with respect to $\mathcal{G}_{\cdot,m}$ means a stopping time with respect to the first index, i.e., $\{\tau \leq j\} \in \mathcal{G}_{-j,m}$ for all $j \geq 0$. In this case, the σ-field $\mathcal{G}_{-\tau,m}$ contains all events A such that $A \cap \{\tau \leq j\} \in \mathcal{G}_{-j,m}$, for all $j \geq 0$.

Definition 4.1 (verifiable time). An a.s. finite nonnegative random time β is said to be verifiable with respect to an increasing family of σ-fields $\{\mathcal{G}_{-j,m}, -j \leq 0 \leq m\}$, if there exists a sequence of random times $\{\beta(m), m \geq 0\}$, with $\beta(m)$ being a backwards $\mathcal{G}_{\cdot,m}$-stopping time for all m, such that

(i) $\beta = \sup_{m \geq 0} \beta(m)$;

(ii) for all $m \geq 0$, $X_m^{-n} = X_m^{-(n+i)}$ for all $i \geq 0$, a.s. on $\{n \geq \beta(m)\}$;

(iii) for all $m \geq 0$, the random variable $X_m^{-\beta(m)}$ is $\mathcal{G}_{-\beta(m),m}$-measurable.

Some comments: First, observe that if β is any backwards coupling time, then it is always possible to find $\beta(m)$ such that (i) and (ii) above hold. The additional thing here is that the $\beta(m)$ are backwards stopping times with respect to some σ-fields, and condition (iii). Second, observe that any verifiable time is a backwards coupling time. This follows directly from (i), (ii) and Definition 2.3. Third, define

$$\beta_m = \max(\beta(0), \ldots, \beta(m)),$$

and observe that

$$(X_0^{-t}, \ldots, X_m^{-t}) = (X_0^{-t-1}, \ldots, X_m^{-t-1}) = \cdots, \quad \text{a.s. on } \{t \geq \beta_m\}.$$

Thus, a.s. on $\{t \geq \beta_m\}$, the sequence $(X_0^{-t}, \ldots, X_m^{-t})$ does not change with t. Since it also converges, in total variation, to $(\tilde{X}_0, \ldots, \tilde{X}_m)$, where $\tilde{X}$ is the stationary version of X, it follows that

$$(X_0^{-t}, \ldots, X_m^{-t}) = (\tilde{X}_0, \ldots, \tilde{X}_m), \quad \text{a.s. on } \{t \geq \beta_m\}.$$

Therefore,

$$(X_0^{-\beta_m}, \ldots, X_m^{-\beta_m}) = (\tilde{X}_0, \ldots, \tilde{X}_m), \quad \text{a.s.}$$

Since $\beta_m \geq \beta(i)$, for each $0 \leq i \leq m$, we have $X_i^{-\beta_m} = X_i^{-\beta(i)}$, and this is $\mathcal{G}_{-\beta(i),i}$-measurable and so, a fortiori, $\mathcal{G}_{-\beta_m,m}$-measurable (the σ-fields are increasing). Thus, $(\tilde{X}_0, \ldots, \tilde{X}_m)$ is $\mathcal{G}_{-\beta_m,m}$-measurable. In other words, any finite-dimensional projection $(\tilde{X}_0, \ldots, \tilde{X}_m)$ of the stationary distribution can be "perfectly sampled". That is, in practice, $\{\mathcal{G}_{-j,m}\}$ contains our basic data (e.g., it measures the random numbers we are using), β_m is a stopping time, and $(\tilde{X}_0, \dots, \tilde{X}_m)$ is measurable with respect to a stopped σ-field. This is what perfect sampling means, in an abstract setup, *without* reference to any Markovian structure.

Naturally, we would like to have a condition for verifiability. Here we present a sufficient condition for the case where renovating events of special structure exist. To prepare for the theorem below, consider a stochastic process $\{X_n, n \in \mathbf{Z}\}$ on $(\Omega, \mathcal{F}, \mathsf{P}, \theta)$, the notation being that of Section 2. Let $\{\zeta_n = \zeta_0 \circ \theta^n, n \in \mathbf{Z}\}$ be a family of i.i.d. random variables. For fixed $\kappa \in \mathbf{Z}$, consider the increasing family of σ-fields

$$\mathcal{G}_{-j,m} := \sigma(\zeta_{-j-\kappa}, \dots, \zeta_m).$$

Consider also a family $\{B_n, n \in \mathbf{Z}\}$ of Borel sets and introduce the events

$$\begin{aligned} A_{-j,m} &:= \{\zeta_{-j-\kappa} \in B_{-\kappa}, \dots, \zeta_m \in B_{m+j}\}, \\ A_0 &:= \bigcap_{m \geq 0} A_{0,m} = \{\zeta_{-\kappa} \in B_{-\kappa}, \dots, \zeta_0 \in B_0, \dots\}, \\ A_n &:= \{\zeta_{n-\kappa} \in B_{-\kappa}, \dots, \zeta_n \in B_0, \dots\} = \theta^{-n} A_0. \end{aligned}$$

Theorem 4.1 (verifiability criterion). *With the notation just introduced, suppose* $\mathsf{P}(A_0) > 0$*. Suppose the* A_n *are renovating events for the process* X*, in the sense that the assumptions of Theorem* 3.1 *hold, and that* $X_m^{-i} \mathbf{1}_{A_{-j,m}}$ *is* $\mathcal{G}_{-j,m}$*-measurable, for all* $-i \leq -j \leq m$*. Then*

$$\beta := \inf\{n \geq 0 : \ \mathbf{1}_{A_{-n}} = 1\}$$

is a verifiable time with respect to the $\{\mathcal{G}_{-j,m}\}$*.*

Proof. We shall show that $\beta = \sup_{m \geq 0} \beta(m)$, for appropriately defined backwards $\mathcal{G}_{\cdot,m}$-stopping times $\beta(m)$ that satisfy the properties (i), (ii) and (iii) of Definition 4.1. Let

$$\beta(m) := \inf\{j \geq 0 : \ \mathbf{1}_{A_{-j,m}} = 1\}.$$

Since $A_{-j,m} \in \mathcal{G}_{-j,m}$, we immediately have that $\beta(m)$ is a backwards $\mathcal{G}_{\cdot,m}$-stopping time. Then

$$\beta(m) := \inf\{j \geq 0 : \ \zeta_{-j-\kappa} \in B_{-\kappa}, \dots, \zeta_m \in B_{m+j}\}$$

is a.s. increasing in m, with

$$\sup_m \beta(m) := \inf\{j \geq 0 : \ \zeta_{-j-\kappa} \in B_{-\kappa}, \dots\} = \inf\{j \geq 0 : \ \mathbf{1}_{A_{-j}} = 1\} = \beta.$$

Hence (i) of Definition 4.1 holds. We next use the fact that the A_n are renovating events. As in (3.5) of the proof of Theorem 3.1 we have, for all $i \geq j$,

$$X_m^{-i} \mathbf{1}_{A_{-j}} = X_m^{-j} \mathbf{1}_{A_{-j}}, \quad \text{a.s.}$$

Since
$$A_{-j} = A_{-j,m} \cap \{\zeta_{m+1} \in B_{m+1+j}, \ldots\} =: A_{-j,m} \cap D_{j,m},$$
we have
$$X_m^{-i} \mathbf{1}_{A_{-j,m}} \mathbf{1}_{D_{j,m}} = X_m^{-j} \mathbf{1}_{A_{-j,m}} \mathbf{1}_{D_{j,m}}, \quad \text{a.s.}$$
By assumption, $X_m^{-i} \mathbf{1}_{A_{-j,m}}$ is $\mathcal{G}_{-j,m}$-measurable, for all $i \geq j$. By the independence between the ζ_n's, $D_{j,m}$ is independent of $\mathcal{G}_{-j,m}$. Hence, by Lemma A.1 of the Appendix, we can cancel the $\mathbf{1}_{D_{j,m}}$ terms in the above equation to get
$$X_m^{-i} \mathbf{1}_{A_{-j,m}} = X_m^{-j} \mathbf{1}_{A_{-j,m}}, \quad \text{a.s.},$$
for all $i \geq j$. Now,
$$\{\beta(m) = j\} \subseteq A_{-j,m}, \tag{4.1}$$
and so, by multiplying by $\mathbf{1}(\beta(m) = j)$ both sides, we obtain
$$X_m^{-i} \mathbf{1}(\beta(m) = j) = X_m^{-j} \mathbf{1}(\beta(m) = j), \quad \text{a.s.},$$
for all $i \geq j$. By Theorem 3.1, $\beta(m) < \infty$, a.s., and so for all $\ell \geq 0$,
$$X_m^{-\beta(m)-\ell} = X_m^{-\beta(m)}, \qquad \text{a.s.}$$
Hence Definition 4.1, (ii) holds. Finally, to show that $X_m^{-\beta(m)}$ is $\mathcal{G}_{-\beta(m),m}$-measurable, we show that $X_m^{-j} \mathbf{1}(\beta(m) = j)$ is $\mathcal{G}_{-j,m}$-measurable. Using the inclusion (4.1) again, we write
$$X_m^{-j} \mathbf{1}(\beta(m) = j) = X_m^{-j} \mathbf{1}_{A_{-j,m}} \mathbf{1}(\beta(m) = j).$$
By assumption, $X_m^{-j} \mathbf{1}_{A_{-j,m}}$ is $\mathcal{G}_{-j,m}$-measurable, and so is $\mathbf{1}(\beta(m) = j)$. Hence Definition 4.1, (iii) also holds. □

A perfect simulation algorithm

In the remaining of this section, we describe a "perfect simulation algorithm", i.e., a method for drawing samples from the stationary version of a process. The setup is as in Theorem 4.1. For simplicity, we take $\kappa = 0$. That is, we assume that
$$A_0 = \{\zeta_0 \in B_0, \zeta_1 \in B_1, \ldots\}$$
has positive probability, and that the $A_n = \theta^{-n} A_0$ are renovating events for the process $\{X_n\}$. Recall that the $\{\zeta_n = \zeta_0 \circ \theta^n\}$ are i.i.d., and that $\mathcal{G}_{m,n} = \sigma(\zeta_m, \ldots, \zeta_n)$, $m \leq n$. It was proved in Theorem 4.1 that the time $\beta = \inf\{n \geq 0 : \ \mathbf{1}_{A_{-n}} = 1\}$ is a bc-time which is verifiable with respect to the $\{\mathcal{G}_{m,n}\}$. This time is written as $\beta = \sup_{m \geq 0} \beta(m)$, where $\beta(m) = \inf\{j \geq 0 : \ \zeta_{-j} \in B_0, \ldots, \zeta_m \in B_{m+j}\}$. The algorithm uses $\beta(0)$ only. It is convenient to let
$$\begin{aligned} \nu_1 &:= \beta(0) = \inf\{j \geq 0 : \ \zeta_{-j} \in B_0, \ldots, \zeta_0 \in B_j\}, \\ \nu_{i+1} &:= \nu_i + \beta(0) \circ \theta^{-\nu_i}, \quad i \geq 1. \end{aligned}$$

In addition to the above, we are going to assume that

$$B_0 \subseteq B_1 \subseteq B_2 \subseteq \dots$$

It is easy to see that this monotonicity assumption is responsible for the following (see Figure 1 for an illustration)

$$\nu_1 \circ \theta^{-j} \leq \nu_1 - j, \quad \text{a.s. on } \{\nu_1 \geq j\}. \tag{4.2}$$

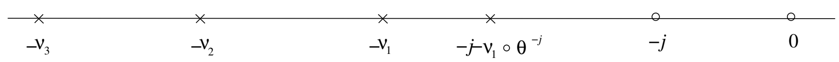

Figure 1. Illustration of the bc-times involved in the perfect simulation algorithm for the stationary process at negative times.

Owing to condition (ii) of Definition 4.1 we have

$$X_0^{-\nu_1} = \tilde{X}_0 = X_0^{-\nu_1 - i}, \quad \forall i \geq 0.$$

That is, if we "start" the process at time $-\nu_1$, we have, at time 0, that X_0 is a.s. equal to the stationary $\tilde{X}_0$. Applying θ^{-j} at this equality we have $X_0^{-\nu_1} \circ \theta^{-j} = \tilde{X}_0 \circ \theta^{-j} = \tilde{X}_{-j}$. But $X_0^{-\nu_1} \circ \theta^{-j} = (X_{\nu_1} \circ \theta^{-\nu_1}) \circ \theta^{-j} = X_{-\nu_1 \circ \theta^{-j}} \circ \theta^{-\nu_1 \circ \theta^{-j} - j} = X_{-j}^{-j-\nu_1 \circ \theta^{-j}}$. That is,

$$X_{-j}^{-j-\nu_1 \circ \theta^{-j}} = \tilde{X}_{-j} = X_{-j}^{-j-\nu_1 \circ \theta^{-j} - i}, \quad \forall i \geq 0. \tag{4.3}$$

But from (4.2), we have $\nu_1 \geq j + \nu_1 \circ \theta^{-j}$, if $\nu_1 \geq j$, and so, from (4.3),

$$X_{-j}^{-\nu_1} = \tilde{X}_{-j}, \qquad \text{a.s. on } \{\nu_1 \geq j\}.$$

This means that if we start the process at $-\nu_1$, then its values on any window $[-j, 0]$ contained in $[-\nu_1, 0]$ match the values of its stationary version on the same window

$$(X_{-j}^{-\nu_1}, \dots, X_0^{-\nu_1}) = (\tilde{X}_{-j}, \dots, \tilde{X}_0), \qquad \text{a.s. on } \{\nu_1 \geq j\}. \tag{4.4}$$

It remains to show a measurability property of the vector (4.4) that we are simulating. By (iii) of Definition 4.1, we have that $X_0^{-\nu_1}$ is $\mathcal{G}_{-\nu_1,0}$-measurable. That is, if $\nu_1 = \ell$ then $\tilde{X}_0$ is a certain deterministic function of $\zeta_{-\ell}, \dots, \zeta_0$. Thus, the functions h_ℓ are defined, for all $\ell \geq 0$, by the condition

$$X_0^{-\ell} = h_\ell(\zeta_{-\ell}, \dots, \zeta_0), \qquad \text{a.s. on } \{\nu_1 = \ell\},$$

or,

$$X_0^{-\nu_1} = h_{\nu_1}(\zeta_{-\nu_1}, \ldots, \zeta_0).$$

Hence for any $i \geq 0$,

$$X_{-i}^{-i-\nu_1\circ\theta^{-i}} = X_0^{-\nu_1}\circ\theta^{-i} = h_{\nu_1\circ\theta^{-i}}(\zeta_{-i-\nu_1\circ\theta^{-i}}, \ldots, \zeta_{-i}).$$

But if $\nu_1 \geq j$, we have $\nu_1\circ\theta^{-i} \leq \nu_1 - i$ for all $i \in [0, j]$, and so every component of $(X_0^{-\nu_1}\circ\theta^{-i}, 0 \leq i \leq j)$ is a deterministic function of $\zeta_0, \ldots, \zeta_{-\nu_1}$. Thus the vector appearing in (4.4) is a deterministic function of $\zeta_0, \ldots, \zeta_{-\nu_1}$, if $\nu_1 \geq j$. This is precisely the measurability property we need.

We now observe that, in (4.4), we can replace ν_1 by any ν_i

$$(X_{-j}^{-\nu_i}, \ldots, X_0^{-\nu_i}) = (\tilde{X}_{-j}, \ldots, \tilde{X}_0), \quad \text{a.s. on } \{\nu_i \geq j\}, \quad i = 1, 2, \ldots$$

Hence if we want to simulate $(\tilde{X}_{-j}, \ldots, \tilde{X}_0)$ we search for an i such that $\nu_i \geq j$, and start the process from $-\nu_i$. It is now clear how to simulate the process on any window prior to 0.

To proceed forward, i.e., to simulate $\{\tilde{X}_n, n > 0\}$, consider first $\tilde{X}_1$. Note that

$$\tilde{X}_1 = \tilde{X}_0\circ\theta = h_{\nu_1}(\zeta_{-\nu_1}, \ldots, \zeta_0)\circ\theta = h_{\nu_1\circ\theta}(\zeta_{-\nu_1\circ\theta+1}, \ldots, \zeta_1).$$

Next note that $\nu_1\circ\theta$ is either equal to 0, or to $\nu_1 + 1$, or to $\nu_2 + 1 = \nu_1 + \nu_1\circ\theta^{-\nu_1} + 1$, etc. This follows from the definition of ν_1 and ν_i, as well as the monotonicity between the B_j. If $\nu_1 = 0$ (which is to say, $\zeta_1 \in B_0$), then $\tilde{X}_1 = h_0(\zeta_1)$. Otherwise, if $\zeta_1 \notin B_0$, but $\zeta_1 \in B_{\nu_1+1}$, then $\nu_1\circ\theta = \nu_1 + 1$, and so $\tilde{X}_1 = h_{\nu_1+1}(\zeta_{-\nu_1}, \ldots, \zeta_1)$. Thus, for some finite (but random) j (defined from $\zeta_1 \in B_{\nu_j+1} \setminus B_{\nu_j}$), we have $\tilde{X}_1 = h_{\nu_j+1}(\zeta_{-\nu_j}, \ldots, \zeta_1)$. The algorithm proceeds similarly for $n > 1$.

The connection between perfect simulation and backward coupling was first studied by Foss and Tweedie [17].

Weak verifiability

Suppose now that we drop the condition that $\mathsf{P}(A_0) > 0$, but only assume that

$$\beta(0) < \infty, \text{ a.s.}$$

Of course, this implies that $\beta(m) < \infty$, a.s., for all m. Here we can no longer assert that we have sc-convergence to a stationary version, but we can only assert existence in the sense described in the sequel. Indeed, simply the a.s. finiteness of $\beta(0)$ (and not of β) makes the perfect simulation algorithm described above realizable. The algorithm is shift-invariant, hence the process defined by it is stationary. One may call this process a stationary version of X. This becomes precise if $\{X_n\}$ itself is a stochastic recursive sequence, in the sense that the stationary process defined by the algorithm is also a stochastic recursive sequence with the same driver. (See Section 5.)

The construction of a stationary version, under the weaker hypothesis $\beta(0) < \infty$, a.s., is also studied by Comets et al. [14], for a particular model. In that paper, it is shown that $\beta(0) < \infty$ a.s., iff

$$\sum_{n=1}^{\infty} \prod_{k=0}^{n} \mathsf{P}(\zeta_0 \in B_k) = \infty.$$

The latter condition is clearly weaker than $\mathsf{P}(A_0) > 0$. In [14] it is shown that it is equivalent to the non-positive recurrence of a certain Markov chain, a realization which leads directly to the proof of this condition.

5. Strong coupling for stochastic recursive sequences

As in the previous section, let $(\Omega, \mathcal{F}, \mathsf{P}, \theta)$ be a probability space with a P-preserving ergodic flow θ. Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$, $(\mathcal{Y}, \mathcal{B}_{\mathcal{Y}})$ be two measurable spaces. Let $\{\xi_n, n \in \mathbf{Z}\}$ be a stationary sequence of $\mathcal{Y}$-valued random variables. Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$ be a measurable function. A stochastic recursive sequence (SRS) $\{X_n, n \geq 0\}$ is defined as an $\mathcal{X}$-valued process that satisfies

$$X_{n+1} = f(X_n, \xi_n), \quad n \geq 0. \tag{5.1}$$

The pair $(f, \{\xi_n\})$ is referred to as the driver of the SRS X. The choice of 0 as the starting point is arbitrary.

A *stationary solution* $\{\tilde{X}_n\}$ of the stochastic recursion is a stationary sequence that satisfies the above recursion. Clearly, it can be assumed that $\tilde{X}_n$ is defined for all $n \in \mathbf{Z}$. There are examples that show that a stationary solution may exist but may not be unique. The classical such example is that of a two-server queue, which satisfies the so-called Kiefer – Wolfowitz recursion (see [11]). In this example, under natural stability conditions, there are infinitely many stationary solutions, one of which is "minimal" and another "maximal". One may define a particular solution, say $\{X_n^0\}$, to the two-server queue SRS by starting from the zero initial condition. Then X^0 sc-converges (under some conditions) to the minimal stationary solution. In our terminology, we may say that the minimal stationary solution is the stationary version of X^0.

Stochastic recursive sequences are ubiquitous in applied probability modeling. For instance, a Markov chain with values in a countably generated measurable space can be expressed in the form of SRS with i.i.d. drivers.

The previous notions of coupling take a simpler form when stochastic recursive sequences are involved owing to the fact that if two SRS with the same driver agree at some n, then they agree thereafter. We thus have the following modifications of the earlier theorems.

Proposition 5.1. *Let X, $\tilde{X}$ be SRS with the same driver $(f, \{\xi_n\})$, and assume that $\tilde{X}$ is stationary. Then*

(i) *X c-converges to $\tilde{X}$ iff* $\lim_{n\to\infty} \mathsf{P}(X_n = \tilde{X}_n) = 1$;

(ii) *X sc-converges to $\tilde{X}$ iff* $\lim_{n\to\infty} \mathsf{P}(\tilde{X}_n = X_n = X_n^{-1} = X_n^{-2} = \cdots) = 1$;

(iii) *X bc-converges iff* $\lim_{n\to\infty} \mathsf{P}(X_0^{-n} = X_0^{-(n+1)} = X_0^{-(n+2)} = \cdots) = 1$.

The standard renovation theory (see [7,8]) is formulated as follows. Our goal here is to show that it fits within the general framework of extended renovation theory, developed in the previous section. First, define renovation events.

Definition 5.1 (renovation event for SRS). Fix $n \in \mathbf{Z}$, $\ell \in \mathbf{Z}_+$ and a measurable function $g : \mathcal{Y}^{\ell+1} \to \mathcal{X}$. A set $R \in \mathcal{F}$ is called (n, ℓ, g)-renovating for the SRS X iff

$$X_{n+\ell+1} = g(\xi_n, \dots, \xi_{n+\ell}), \quad \text{a.s. on } R. \tag{5.2}$$

An alternative terminology is: R is a renovation event on the segment $[n, n+\ell]$ (see [8]).

We then have the following theorem, which is a special case of Theorem 3.1.

Theorem 5.1 (renovation theorem for SRS). *Fix $\ell \geq 0$ and $g : \mathcal{Y}^{\ell+1} \to \mathcal{X}$. Suppose that, for each $n \geq 0$, there exists a (n, ℓ, g)-renovating event R_n for X. Assume that $\{R_n, n \geq 0\}$ is stationary and ergodic, with $\mathsf{P}(R_0) > 0$. Then the SRS X bc-converges and its stationary version $\tilde{X}$ is an SRS with the same driver as X.*

Proof. For each $n \in \mathbf{Z}$, define $\hat{X}_{n,i}$, recursively on the index i, by

$$\begin{aligned} \hat{X}_{n,n+\ell+1} &= g(\xi_n, \dots, \xi_{n+\ell}), \\ \hat{X}_{n,n+j+1} &= f(\hat{X}_{n,n+j}, \xi_{n+j}), \quad j \geq \ell + 1, \end{aligned} \tag{5.3}$$

and observe that (5.2) implies that

$$\forall n \geq 0, \ \ \forall j \geq \ell + 1, \qquad X_{n+j} = \hat{X}_{n,n+j}, \quad \text{a.s. on } R_n. \tag{5.4}$$

For $n \in \mathbf{Z}$, set $A_n := R_{n-\ell-1}$. Consider the stationary background

$$H_{n,p} := \hat{X}_{n-\ell-1,p}, \quad p \geq n.$$

Note that $H_{n,p} \circ \theta^k = H_{n+k,p+k}$ and rewrite (5.4) as

$$\forall n \geq \ell + 1, \ \ \forall i \geq 0, \qquad X_{n+i} = H_{n,n+i}, \quad \text{a.s. on } A_n,$$

which is precisely condition (3.1) of Theorem 3.1. Since $\mathsf{P}(A_0) = \mathsf{P}(R_0) > 0$, Theorem 3.1 applies and allows us to conclude that there is a unique stationary version $\tilde{X}$, constructed by means of the bc-time

$$\gamma := \inf\{i \geq 0 : \mathbf{1}_{R_{-i-\ell-1}} = 1\}. \tag{5.5}$$

We have

$$\tilde{X_n} = X_{\gamma+n} \circ \theta^{-\gamma} = H_{-\gamma,n} = \hat{X}_{-\gamma-\ell-1,n}.$$

This $\tilde{X}_n$ can be defined for all $n \in \mathbf{Z}$. From this and (5.3), we have that $\tilde{X}_{n+1} = f(\tilde{X}_n, \xi_n)$, $n \in \mathbf{Z}$, i.e., $\tilde{X}$ has the same driver as X. □

It is useful to observe that, if R_n are (n, ℓ, g) renovating events for X with $\mathsf{P}(R_0) > 0$, then the stationary version $\tilde{X}$ satisfies $\tilde{X}_{-\gamma} = g(\xi_{-\gamma-\ell-1}, \ldots, \xi_{-\gamma-1})$, a.s., where γ is the bc-time defined in (5.5). More generally, if we consider the random set $\{j \in \mathbf{Z} : \mathbf{1}_{R_j} = 1\}$ (the set of renovation epochs), we have, for any α in this set, $\tilde{X}_\alpha = g(\xi_{\alpha-\ell-1}, \ldots, \xi_{\alpha-1})$, a.s.

6. Strong coupling for functionals of stochastic recursive sequences

The formulation above extends easily to functionals of stochastic recursions. Namely, suppose that, in addition to the SRS X satisfying (5.1), we also have a process Z with values in a third measurable space $(\mathcal{Z}, \mathcal{B}_\mathcal{Z})$ given by

$$Z_n = \varphi(X_n, \xi_n), \tag{6.1}$$

where $\varphi : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ is a certain measurable function. Given an integer $\ell \geq 0$ and a measurable function $g : \mathcal{Y}^{\ell+1} \to \mathcal{X}$, define, for each $n \in \mathbf{Z}$, the quantities $\hat{X}_{n,i}$ and $\hat{Z}_{n,i}$, recursively in the index i, by

$$\begin{aligned}
\hat{X}_{n,n+\ell+1} &= g(\xi_n, \ldots, \xi_{n+\ell}), \\
\hat{X}_{n,n+j+1} &= f(\hat{X}_{n,n+j}, \xi_{n+j}), \quad j \geq \ell + 1, \\
\hat{Z}_{n,n+j} &= \varphi(\hat{X}_{n,n+j}, \xi_{n+j}), \quad j \geq \ell + 1.
\end{aligned}$$

Definition 6.1 (renovation event for functionals of SRS). A set $R \in \mathcal{F}$ is said to be (n, ℓ, g)-renovating for Z iff

$$\forall j \geq \ell + 1, \qquad Z_{n+j} = \hat{Z}_{n,n+j}, \quad \text{a.s. on } R. \tag{6.2}$$

Theorem 6.1 (renovation theorem for functionals of SRS). *Let us suppose that, for each $n \geq 0$, there exists a (n, ℓ, g)-renovating event R_n for Z. Assume that $\{R_n, n \geq 0\}$ is stationary and ergodic, with $\mathsf{P}(R_0) > 0$. Then Z bc-converges.*

Proof. Let $A_n := R_{n-\ell-1}$, and $H_{n,p} := \hat{Z}_{n-\ell-1,p}$, $p \geq n$. Observe that $H_{n,p} \circ \theta^k = H_{n+k,p+k}$. Then (6.2) is written as

$$\forall n \geq \ell+1, \quad \forall i \geq 0, \qquad Z_{n+i} = H_{n,n+i}, \quad \text{a.s. on } A_n.$$

This is again (3.1) and Theorem 3.1 applies. □

A bc-time is given by the formula $\gamma = \inf\{i \geq 0 : \mathbf{1}_{R_{-i-\ell-1}} = 1\}$. The stationary version $\tilde{Z}$ is given by $\tilde{Z}_n = H_{-\gamma,n} = \hat{Z}_{-\gamma-\ell-1,n}$. In particular, $\tilde{Z}_{-\gamma} = \varphi(g(\xi_{-\gamma-\ell-1}, \ldots, \xi_{-\gamma-1}), \xi_{-\gamma})$. Functionals of SRS are particularly important in applications, as in, e.g., Monte Carlo simulation methods. Although, as before, for any renovation epoch α, i.e., any element of the random set $\{j \in \mathbf{Z} : \mathbf{1}_{R_j} = 1\}$, we have $\tilde{Z}_\alpha = \varphi(g(\xi_{\alpha-\ell-1}, \ldots, \xi_{\alpha-1}), \xi_\alpha)$, it is not clear that these epochs are possible to simulate because they may depend on the entire history of the driver. Thus, the concept of verifiability introduced earlier is particularly relevant here. It is important to define renovation events so that they yield verifiable bc-times. This is frequently possible. We give an example.

Example 6.1. Consider a Markov chain $\{X_n\}$ with values in a finite set S, having stationary transition probabilities $p_{i,j}$, $i, j \in S$. Assume that $[p_{i,j}]$ is irreducible and aperiodic. Although there is a unique invariant probability measure, whether X bc-converges to the stationary Markov chain $\tilde{X}$ depends on the realization of X on a particular probability space. We can achieve bc-convergence with a verifiable bc-time if we realize X as follows: Consider a sequence of i.i.d. random maps $\xi_n : S \to S$, $n \in \mathbf{Z}$ (and we write $\xi_n \xi_{n+1}$ to indicate composition). Represent each ξ_n as a vector $\xi_n = (\xi_n(i), i \in S)$, with *independent* components such that

$$\mathsf{P}(\xi_n(i) = j) = p_{i,j}, \quad i, j \in S.$$

Then the Markov chain is realized as an SRS by

$$X_{n+1} = \xi_n(X_n).$$

It is important to notice that the condition that the components of ξ_n be independent is not necessary for the Markovian property. It is only used as a means of constructing the process on a particular probability space, so that backwards coupling takes place. Now define

$$\beta = \inf\{n \geq 0 : \forall\, i, j \in S, \quad \xi_0 \cdots \xi_{-n}(i) = \xi_0 \cdots \xi_{-n}(j)\}.$$

It can be seen that, under our assumptions, β is a bc-time for X, $\beta < \infty$, a.s., and β is verifiable. This bc-time is the one used by Propp and Wilson [22] in their perfect simulation method for Markov chains. Indeed, the verifiability property of β allows recursive simulation of the random variable $\xi_0 \cdots \xi_{-\beta}(i)$ which (regardless of i) has the stationary distribution.

Example 6.2. Another interesting example is the process considered in [12] by Brémaud and Massoulié. This process has a "Markov-like" property with *random memory*. Consider a process $\{X_n, n \in \mathbf{Z}\}$ with values in a Polish space S and suppose that its transition kernel, defined by $\mathsf{P}(X_n \in \cdot \mid X_{n-1}, X_{n-2}, \ldots)$ is time-homogeneous (a similar setup is considered in [14]), i.e. that there exists a kernel $\mu : K^\infty \times \mathcal{B}(K) \to [0,1]$ such that $\mathsf{P}(X_n \in B \mid X_{n-1}, X_{n-2}, \ldots) = \mu((X_{n-1}, X_{n-2}, \ldots); B)$, $B \in \mathcal{B}(K)$. This represents the dynamics of the process. In addition, assume that the dynamics does not depend on the whole past, but on a finite but random number of random variables from the past. It is also required that the random memory is "consistent" and that the minorization condition holds, i.e., $\mu((X_{n-1}, X_{n-2}, \ldots), \cdot) \geq \varepsilon\nu(\cdot)$, where $\varepsilon \in (0,1)$ and ν is a probability measure on $(K, \mathcal{B}(K))$. See [12] for details. Then it is shown that renovation events do exist and that the process $\{X_n\}$ sc-converges to a stationary process that has the same dynamics μ.

7. The infinite bin model

Consider an infinite number of bins arranged on the line and indexed, say, by the non-positive integers. Each bin can contain an unlimited but finite number of particles. A configuration is a finite-dimensional vector

$$x = [x(-\ell), \ldots, x(-1), x(0)], \tag{7.1}$$

if there are only finitely many particles in the system, or an infinite vector, otherwise. (The unconventional indexing by non-positive integers is done for reasons of convenience when using the infinite bin model to understand asymptotics of stochastic ordered graphs.) At each integer step, precisely one particle of the current configuration is chosen in some random manner. If the particle is in bin $-i \leq -1$, then a new particle is created and placed in bin $-i+1$. Otherwise, if the chosen particle is in bin 0 then a new bin is created to hold the child particle and a relabeling of the bins occurs: the existing ones are shifted by one place to the left (and are re-indexed) and the new bin is given the label 0. This kind of system, besides being interesting in its own right, will later be derived from a stochastic ordered graph.

To make the verbal description above precise, we introduce some notations. Define the configuration space

$$S = \bigcup_{1 \leq n \leq \infty} \mathbf{N}^n = \mathbf{N}^* \cup \mathbf{N}^\infty,$$

where $\mathbf{N} = \{1, 2, \ldots\}$ is the set of natural numbers, $\mathbf{N}^{n+1} = \mathbf{N}^n \times \mathbf{N}$, $n \geq 1$, $\mathbf{N}^* = \bigcup_{1 \leq n < \infty} \mathbf{N}^n$ is the set of all finite-dimensional vectors, and $\mathbf{N}^\infty$ is the set of all infinite-dimensional vectors. We endow S with the natural topology of pointwise convergence, and let $\mathcal{B}_S$ be the corresponding class of Borel sets generated by this topology. The *extent* of an $x \in \mathbf{N}^*$ is defined as

$$|x| = \ell, \quad \text{if } x = [x(-\ell), \ldots, x(0)],$$

and the *norm* as

$$||x|| = \sum_{j=0}^{\ell} x(-j),$$

and if $x \in \mathbf{N}^\infty$, we set $|x| = ||x|| = +\infty$. The *concatenation* of an $x \in S$ with a $y \in \mathbf{N}^*$ is defined by

$$[x, y] = [\ldots, x(-1), x(0), y(-|y|), \ldots, y(0)].$$

The *restriction* $x^{(-k)}$ of $x \in S$ onto $\{-k, \ldots, 0\}$, where $k \le |x|$, is defined by

$$x^{(-k)} = [x(-k), \ldots, x(0)].$$

Finally, for any $x \in S$, and any integer $0 \le k \le |x|$, let $x + \delta_{-k}$ be the vector obtained from x by adding 1 to the coordinate $x(-k)$ and leaving everything else unchanged:

$$x + \delta_{-k} = [\ldots,\ x(-k-1),\ x(-k)+1,\ x(k+1), \ldots, x(-1),\ x(0)].$$

The dynamics of the model requires defining the map $\Phi : S \times \mathbf{N} \to S$ by

$$\Phi(x, \xi) = \begin{cases} [x, 1], & \text{if } \ \xi \le x(0), \\ x + \delta_{-k}, & \text{if } \ \sum_{j=0}^{k} x(-j) < \xi \le \sum_{j=0}^{k+1} x(-j), \quad 0 \le k < |x|, \\ x + \delta_{-|x|}, & \text{if } \ \xi > ||x||. \end{cases} \tag{7.2}$$

Given a stationary-ergodic sequence $\{\xi_n, n \in \mathbf{Z}\}$ of $\mathbf{N}$-valued random variables, and an S-valued random variable X_0, define a stochastic recursive sequence by

$$X_{n+1} = \Phi(X_n, \xi_{n+1}), \quad n \ge 0. \tag{7.3}$$

Our goal is to find conditions under which there is a unique stationary solution to the above stochastic recursion and that any solution strongly converges to it. A special case, namely when the $\{\xi_n\}$ are i.i.d., and, consequently, X is Markovian, will also be considered.

Extensions of the infinite bin model, which can be handled using the techniques developed here, are possible. For example, we can consider the contents of the bins to be positive valued (as opposed to just integer-valued) random variables. These are natural models of random graphs with random weights, and will be considered in a forthcoming paper.

It is clear from the outset that the decomposition of the configuration space into $S = \mathbf{N}^* \cup \mathbf{N}^\infty$ is a decomposition into a 'transient part' ($\mathbf{N}^*$) and a 'non-transient part' ($\mathbf{N}^\infty$). Indeed, regardless of the initial configuration X_0, the extent of X_n grows as n grows. Thus, it is intuitively clear that a stationary solution $\tilde{X}$, if any, will be such that $\mathsf{P}(\tilde{X}_0 \in \mathbf{N}^*) = 0$. This means that it is not possible to construct renovation events for the process $\{X_n\}$ satisfying (7.3). (Otherwise, we would have coupling between the process $\{X_n\}$ that is supported on $\mathbf{N}^*$ and the process $\{\tilde{X}_n\}$ that is supported outside $\mathbf{N}^*$). Instead we consider functionals of it, namely, all its finite-dimensional projections

$$X_n^{(-\ell)} = [X_n(-\ell), \dots, X_n(0)], \quad \ell = 0, 1, 2, \dots,$$

and find renovation events for each of them.

We start with a couple of properties of the map Φ defined by (7.2).

Lemma 7.1. *For any $x \in S$, $y \in \mathbf{N}^*$, if $\xi \le ||y|| + 1$, then*

$$\Phi([x, y], \xi) = [x, \Phi(y, \xi)].$$

Furthermore, if $|y| \ge \ell + 1$, for some $\ell \ge 0$, then, for all $x, x' \in S$, and all $\xi \in \mathbf{N}$,

$$\Phi([x, y], \xi)^{(-\ell)} = \Phi([x', y], \xi)^{(-\ell)}.$$

Proof. If $\xi \le y(0)$, then $\Phi([x, y], \xi) = [x, y, 1]$, and also $\Phi(y, \xi) = [y, 1]$, so the statement is true. Otherwise, if $\xi > y(0)$, $\Phi([x, y], \xi) = [x, y] + \delta_{-k}$ for some k smaller than or equal to $|y|$ because of the condition $\xi \le ||y|| + 1$, as can be verified by the second line of (7.2). But then $\Phi([x, y], \xi) = [x, y] + \delta_{-k} = [x, y + \delta_{-k}]$. Also, again by the assumption $\xi \le ||y|| + 1$, and the second line of (7.2), $\Phi(y, \xi) = y + \delta_{-k}$, for the same k. Hence, here too, $\Phi([x, y], \xi) = [x, y + \delta_{-k}] = [x, \Phi(y, \xi)]$. The second identity follows immediately from (7.2). $\square$

This lemma motivates the consideration of the following events, which will serve as renovation events

$$A_0^\ell = \bigcap_{j=1}^{\ell+1} \{\xi_j = 1\} \cap \bigcap_{j=\ell+2}^{\infty} \{\xi_j \le j\}, \quad \ell \ge 0. \tag{7.4}$$

Lemma 7.2. *Let $\{X_n, n \ge 0\}$ be the stochastic recursive sequence* (7.3) *describing the infinite bin model. Define random variables Y_n^ℓ, $n \ge \ell + 1$, $\ell \ge 0$, by*

$$\begin{aligned} Y_{\ell+1}^\ell &= [1, \dots, 1] \quad \text{(a vector of } \ell + 1 \text{ consecutive 1s)}, \\ Y_{n+1}^\ell &= \Phi(Y_n^\ell, \xi_{n+1}), \quad n \ge \ell + 1. \end{aligned} \tag{7.5}$$

Then, for all $\ell \geq 0$, it holds that

$$\forall\ n \geq \ell+1, \qquad X_n = [X_0, Y_n^\ell], \quad \text{a.s. on } A_0^\ell. \tag{7.6}$$

Furthermore, $|Y_n^\ell| \geq \ell$, $||Y_n^\ell|| = n$, for all $n \geq \ell+1 \geq 1$.

Proof. First observe that the last statement about the extent and norm of Y_n^ℓ follows from the obvious properties of Φ:

$$|\Phi(x,\xi)| \geq |x|, \quad ||\Phi(x,\xi)|| = ||\Phi(x,\xi)|| + 1.$$

Next, since, on A_0^ℓ, $\xi_1 = \cdots = \xi_{\ell+1} = 1$, the configuration starts from X_0 and evolves up to $n = \ell+1$ by concatenating a 1 to the right of the previous configuration, namely,

$$X_1 = [X_0, 1],\ X_2 = [X_0, 1, 1], \ldots,\ X_{\ell+1} = [X_0, Y_{\ell+1}^\ell],$$

where $Y_{\ell+1}^\ell$ is a vector of $\ell+1$ consecutive 1s. So the result is true for $n = \ell+1$. We proceed by induction. Our induction hypothesis is that, for some fixed $n \geq \ell+1$, $X_n = [X_0, Y_n^\ell]$, a.s. on A_0^ℓ. We have

$$X_{n+1} = \Phi(X_n, \xi_{n+1}) = \Phi([X_0, Y_n^\ell], \xi_{n+1}).$$

But $||Y_n^\ell|| = n$ and, a.s. on A_0^ℓ, $\xi_{n+1} \leq n+1$. Thus, by the first identity of Lemma 7.1,

$$\Phi([X_0, Y_n^\ell], \xi_{n+1}) = [X_0, \Phi(Y_n^\ell, \xi_{n+1})] = [X_0, Y_{n+1}^\ell],$$

and so $X_{n+1} = [X_0, Y_{n+1}^\ell]$, a.s. on A_0^ℓ. □

We have that $|Y_n^\ell| \geq \ell$ for all $n \geq \ell+1$, and we just showed that $X_n = [X_0, Y_n^\ell]$, a.s. on A_0^ℓ. Thus, a.s. on the same event, the restriction X onto $\{-\ell, \ldots, 0\}$ is the restriction of Y_n^ℓ. In other words, $X_n^{(-\ell)}$ is a function of Y_n^ℓ for all $n \geq \ell+1$, a.s. on A_0^ℓ. Hence A_0^ℓ is a renovation event for $X^{(-\ell)}$, in the sense of Definition 6.1. In fact, since there is nothing special with 0 as the origin of time, we can define, for all s,

$$A_s^\ell := \theta^{-s} A_0^\ell = \bigcap_{j=1}^{\ell+1} \{\xi_{s+j} = 1\} \cap \bigcap_{j=\ell+2}^{\infty} \{\xi_{s+j} \leq j\}. \tag{7.7}$$

Then we have

$$X_{s+\ell+1} = [X_s, Y_{\ell+1}^\ell \circ \theta^s],\ X_{s+\ell+2} = [X_s, Y_{\ell+2}^\ell \circ \theta^s], \ldots, \quad \text{a.s. on } A_s^\ell. \tag{7.8}$$

(Equation (7.6) is just the special case with $s = 0$.) It follows again that, a.s. on A_s^ℓ, and for all $j \geq \ell_1$, the restriction of X_{s+j} onto $\{-\ell, \ldots, 0\}$ is obtained by the restriction of $Y_j^\ell \circ \theta^s$ onto the same set

$$X_{s+\ell+1}^{(-\ell)} = (Y_{\ell+1}^\ell \circ \theta^s)^{(-\ell)}, \quad X_{s+\ell+2}^{(-\ell)} = (Y_{\ell+2}^\ell \circ \theta^s)^{(-\ell)}, \ldots, \quad \text{a.s. on } A_s^\ell.$$

Use the abbreviation

$$H^\ell_{t,u} = (Y^\ell_{\ell+1+(u-t)} \circ \theta^{t-\ell-1})^{(-\ell)}, \quad t \le u, \tag{7.9}$$

(which will be the stationary background process) to summarize the last observations in

Lemma 7.3. *Let A^ℓ_s be defined by* (7.7) *and $H^\ell_{t,u}$ by* (7.9). *Then, for all $\ell \ge 0$,*

$$\forall t \ge \ell+1, \qquad X^{(-\ell)}_t = H^\ell_{t,t},\ X^{(-\ell)}_{t+1} = H^\ell_{t,t+1}, \dots \qquad \text{a.s. on } A^\ell_{t-\ell-1}. \tag{7.10}$$

Furthermore, $H_{t,u} \circ \theta^s = H_{t+s,u+s}$.

Proof. The last relation follows from the fact that Y^ℓ_n is a function of $(\xi_i, \ell+2 \le i \le n)$ (see (7.5)) and (7.9). Finally, (7.10) follows from (7.8) and the fact that the extent of $Y^\ell_{t+i} \circ \theta^t$ is at least ℓ. □

We introduce now a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a measurable flow θ that preserves P. Also, θ is assumed to be ergodic. A random sequence $\{\xi_n, n \in \mathbf{Z}\}$ is then stationary-ergodic if $\xi_n \circ \theta^m = \xi_{n+m}$ for all $m, n \in \mathbf{Z}$. The theorem that follows constitutes the main result of this section.

Theorem 7.1. *Let $\{\xi_n, n \in \mathbf{Z}\}$ be a stationary-ergodic sequence of nontrivial $\mathbf{N}$-valued random variables such that $\mathsf{P}(A^\ell_0) > 0$, for all $\ell \ge 0$, where A^ℓ_0 is defined in* (7.4). *Then there exists a unique stationary and ergodic stochastic recursive sequence $\{\tilde{X}_n, n \in \mathbf{Z}\}$, with values in $\mathbf{N}^\infty$, such that $\tilde{X}_{n+1} = \Phi(\tilde{X}_n, \xi_{n+1})$, for all $n \in \mathbf{Z}$. Furthermore, if X is any SRS satisfying* (7.3), *then the distribution of X_n converges weakly, as $n \to \infty$, to that of $\tilde{X}_0$.*

Proof. Consider the infinite bin model $\{X_n\}$ with arbitrary initial configuration $X_0 \in S$, a.s. Fix $\ell \ge 0$ and consider the process $\{X^{(-\ell)}_n\}$. Observe that the premises of Theorem 3.1 are satisfied, namely, $\{A^\ell_n\}$ is a stationary sequence with $\mathsf{P}(A^\ell_0) > 0$ and $\{H^\ell_{t,u},\ t \le u\}$ is a stationary background process for which the "renovation condition" (7.9) holds. Therefore, by the result of Theorem 3.1 the process $X^{(-\ell)}_n$ bc-converges. Let $\{Z_{\ell,n}, n \in \mathbf{Z}\}$ denote its stationary version. Clearly, this is compatible with the stationary version $\{Z_{\ell+1,n}, n \in \mathbf{Z}\}$ of $X^{(-\ell-1)}_n$, in the sense that $Z_{\ell,n} = Z^{(-\ell)}_{\ell+1,n}$ (the restriction of the latter onto the set $\{-\ell, \dots, 0\}$ is the former). Hence we define uniquely a stationary process $\tilde{X}_n$, with values in $\mathbf{N}^\infty$, by defining $\tilde{X}^{(-\ell)}_n \equiv Z_{\ell,n}$, for all $\ell \ge 0$. To show that $\tilde{X}_n$ satisfies the recursion, it is enough to show that $\tilde{X}^{(-\ell)}_n = \Phi(\tilde{X}_n, \xi_{n+1})^{(-\ell)}$ for all $\ell \ge 0$, and all n. Fix $\ell \ge 0$, and let σ be the minimal coupling time between $X^{(-\ell-1)}_n$ and $\tilde{X}_n$. Then

$$X^{(-\ell-1)}_n = \tilde{X}^{(-\ell-1)}_n, \quad \text{a.s. on } \{n \ge \sigma\}, \tag{7.11}$$

and, a fortiori,

$$X_n^{(-\ell)} = \tilde{X}_n^{(-\ell)}, \quad \text{a.s. on } \{n \ge \sigma\}. \tag{7.12}$$

We then have that, a.s. on $\{n \ge \sigma\}$,

$$\begin{aligned}\tilde{X}_{n+1}^{(-\ell)} = X_{n+1}^{(-\ell)} = \Phi(X_n, \xi_{n+1})^{(-\ell)} &= \Phi\big(\big[\dots, X_n(-\ell-2), X_n^{(-\ell-1)}\big], \xi_{n+1}\big)^{(-\ell)} \\ &= \Phi\big(\big[\dots, X_n(-\ell-2), \tilde{X}_n^{(-(\ell+1))}\big], \xi_{n+1}\big)^{(-\ell)} \\ &= \Phi\big(\big[\dots, \tilde{X}_n(-\ell-2), \tilde{X}_n^{(-\ell-1)}\big], \xi_{n+1}\big)^{(-\ell)} \\ &= \Phi(\tilde{X}_n, \xi_{n+1})^{(-\ell)},\end{aligned}$$

where the first equality is (7.12), the second is the dynamics of X, the third is a rewriting of the second, the fourth is (7.11), the fifth follows from the second identity of Lemma 7.1, and the sixth is a rewriting of the fifth. Since $\tilde{X}$ is stationary, and σ an a.s. finite time, it follows that the proven identity is true a.s. for all n. Finally, since $X_n^{(-\ell)}$ converges weakly, as $n \to \infty$, to $Z_{\ell,0} = \tilde{X}_0^{(-\ell)}$, for each $\ell \ge 0$, we immediately get that X_n converges weakly to $\tilde{X}_0$. □

Some remarks

Fix $\ell \ge 0$ and define the random indices j for which $\mathbf{1}_{A^\ell_{j-\ell-1}} = 1$:

$$D^\ell = \{j \in \mathbf{Z} : \ \mathbf{1}_{A^\ell_{j-\ell-1}} = 1\}.$$

The elements of D^ℓ are called *ℓth order decoupling epochs* and are enumerated according to the convention

$$\cdots < \gamma^\ell_{-1} < \gamma^\ell_0 \le 0 < \gamma^\ell_1 < \gamma^\ell_2 < \cdots$$

We mention that a renovation epoch occurs $\ell + 1$ before a decoupling epoch. We have

$$X_{\gamma^\ell_r} = [X_{\gamma^\ell_r - \ell - 1}, 1, \dots, 1]$$

(i.e., the previous value concatenated with ℓ+1 consecutive 1s.) Every projection of the stationary solution can be constructed starting from any of this points. That is, pick a point $\gamma^\ell_r \in D^\ell$ and define

$$\tilde{X}_{\gamma^\ell_r}^{(-\ell)} = [1, \dots, 1], \qquad (\ \ell + 1 \text{ consecutive 1s}).$$

Then iterate forward according to

$$\tilde{X}_{n+1}^{(-\ell)} = \Phi\big(\tilde{X}_n^{(-\ell)}, \xi_{n+1}\big).$$

In this manner, the sequence $\{\tilde{X}_{\gamma^\ell_r + i}, i \ge 0\}$ is constructed. Note that at the next point γ^ℓ_{r+1} of the point process D^ℓ, it will again be the case that $\tilde{X}_{\gamma^\ell_{r+1}}^{(-\ell)} =$

$[1, \ldots, 1]$, a vector of $\ell + 1$ consecutive 1s; this happens automatically by the renovation that takes place at time $\gamma_{r+1}^{\ell} - \ell - 1$. The point process D^{ℓ} becomes rarer and rarer as ℓ increases, with rate converging to zero (except in the trivial case $\xi_n \equiv 1$, a.s., which is excluded) and thus it is impossible to have the whole process X couple with $\tilde{X}$ at any of those random times. In other words, it is impossible to construct the whole $\tilde{X}$ by this coupling method; only its finite-dimensional projections can be constructed by coupling.

That the condition of Theorem 7.1 is non-vacuous is considered next, in an important special case.

Corollary 7.1. *Suppose* $\{\xi_n, n \in \mathbf{Z}\}$ *are i.i.d. Then the condition* $\mathsf{P}(A_0^{\ell}) > 0$ *for all* $\ell \geq 0$ *is equivalent to*

(i) $\mathsf{P}(\xi_0 = 1) > 0$,

and

(ii) $\mathsf{E}\xi_0 < \infty$.

In this case, the infinite bin model forms is a Markov process.

Proof. We have

$$\mathsf{P}(A_0^{\ell}) = \mathsf{P}(\xi_0 = 1)^{\ell+1} \prod_{k \geq \ell+2} \mathsf{P}(\xi_0 \leq k).$$

Let $G(k) = \mathsf{P}(\xi_0 > k)$ and write

$$\log \mathsf{P}(A_0^{\ell}) = (\ell + 1) \log \mathsf{P}(\xi_0 = 1) + \sum_{k \geq \ell+2} \log(1 - G(k)).$$

Thus, $\log \mathsf{P}(A_0^{\ell}) > -\infty$ implies and is implied by $\mathsf{P}(\xi_0 = 1) > 0$ and $\sum_{k \geq \ell+2} G(k) < \infty$. The last inequality is equivalent to $\mathsf{E}\xi_0 < \infty$. □

8. Functional law of large numbers for the infinite bin model

We now study the growth of partial sums

$$S_n(k) := X_n(-k) + \cdots + X_n(-1) + X_n(0),$$

as $n \to \infty$, in two cases: for an infinite bin model that starts with $X_0 \in \mathbf{N}^*$ (i.e., X_0 has finite extent), and for the stationary infinite bin model. For simplicity, in the first case we will assume X_0 is the trivial configuration with just one particle, $X_0 = [1]$, and call it the *transient infinite bin model*. We will avoid using tildes to denote the stationary version, unless there is a fear of confusion. Throughout this section, we assume only part of the assumptions of

Theorem 7.1, and consider A_n^ℓ only for $\ell = 0$, so, to simplify notation we omit the superscript ℓ everywhere, and let

$$A_n = A_n^0 = \{\xi_{n+1} = 1, \xi_{n+2} \le 2, \xi_{n+3} \le 3, \ldots\},$$

and assume that

$$\rho := \mathsf{P}(A_0) > 0. \tag{8.1}$$

We also assume that the model is not trivial, viz., $\mathsf{P}(\xi_0 = 1) < 1$, and so ρ is a constant strictly between 0 and 1. Consider the *decoupling epochs*

$$D = \{n \in \mathbf{Z} : \ \mathbf{1}_{A_{n-1}} = 1, \text{ a.s.}\}.$$

Thus if the renovation event A_n occurs, then there is a decoupling at time $n+1$ (in words, a new bin is created, a new particle is placed in it, and, thereafter, all particles are never placed in bins occupied prior to time n). Let γ_r denote the elements of D, enumerated as

$$\cdots < \gamma_{-1} < \gamma_0 \le 0 < \gamma_1 < \gamma_2 < \cdots$$

We regard D as a stationary and ergodic point process on the integers with rate ρ. We define

$$\Gamma(m,n] := \sum_{r \in \mathbb{Z}} \mathbf{1}(m < \gamma_r \le n), \quad m \le n,$$

and consider the corresponding counting process, defined by

$$\Gamma_n := \begin{cases} \Gamma(0,n], & n \ge 0, \\ -\Gamma(n,0], & n \le 0, \end{cases}$$

so, in particular, $\Gamma_0 = \Gamma(\emptyset) = 0$, a.s. By ergodicity,

$$\frac{\Gamma_n}{n} \to \rho, \quad \text{as } n \to \pm\infty, \quad \frac{\gamma_r}{r} \to \rho^{-1}, \quad \text{as } r \to \pm\infty, \quad \text{a.s.}$$

Consider now the events

$$B_n = \{\xi_{n+1} \le X_n(0)\}.$$

When B_n occurs there is a *shift* at time $n+1$ (i.e., a new bin is created and a new particle is placed in it). By analogy to the above, we consider the set

$$G = \{n \in \mathbf{Z} : \ \mathbf{1}_{B_{n-1}} = 1, \text{ a.s.}\}.$$

Let δ_r denote the elements of G, enumerated as

$$\cdots < \delta_{-1} < \delta_0 \le 0 < \delta_1 < \delta_2 < \cdots$$

Also, define

$$L(m,n] := \sum_{r\in\mathbb{Z}} \mathbf{1}(m < \delta_r \le n), \quad m \le n,$$

and consider the corresponding counting process, defined by

$$L_n := \begin{cases} L(0,n], & n \ge 0, \\ -L(n,0], & n \le 0. \end{cases}$$

The understanding here is that, if we consider the transient infinite bin model, then L is defined on the non-negative integers. Unlike the decoupling point process, L is stationary only for the stationary case. However, due to the results of the previous section, it is asymptotically stationary and has a well-defined rate. Indeed, for $n > 0$,

$$\frac{L_n}{n} = \frac{1}{n}\sum_{j=0}^{n-1} \mathbf{1}_{B_j} = \frac{1}{n}\sum_{j=0}^{n-1} \mathbf{1}(\xi_{j+1} \le X_j(0)).$$

Since $\rho > 0$, the bivariate process $(\xi_n, X_n(0))$ coupling converges to $(\xi_n, \tilde{X}_n(0))$, so we can replace $X_j(0)$ by $\tilde{X}_j(0)$ in the above sum when taking the limit as $n \to \infty$, to conclude that there is a constant

$$C = \mathsf{P}(\xi_1 \le \tilde{X}_0(0)), \tag{8.2}$$

strictly between 0 and 1, such that

$$\frac{L_n}{n} \to C, \quad \text{as } n \to \infty, \quad \text{a.s.},$$

both for the transient and stationary case. For the stationary case, we also have the same limit when $n \to -\infty$. Since $A_n \subseteq B_n$ for all n, we have

$$0 < \rho \le C < 1.$$

We now state the main theorem of this section. Below, $[x]$ denotes the integer part of the real number x, and $t \wedge s = \min(t,s)$.

Theorem 8.1. *Suppose $\rho > 0$. Let $S_n(k) := X_n(-k) + \cdots + X_n(-1) + X_n(0)$. Then, for any $T > 0$, for the stationary infinite bin model we have*

$$\lim_{n\to\infty} \sup_{0\le t\le T} \left| \frac{1}{n} S_n([nCt]) - t) \right| \to 0, \quad \text{a.s.},$$

and for the transient infinite bin model we have

$$\lim_{n\to\infty} \sup_{0\le t\le T} \left| \frac{1}{n} S_n([nCt]) - (t \wedge 1) \right| \to 0, \quad \text{a.s.}$$

Note that, in the expression for $S_n(k)$, the index k may exceed the extent of X_n, in which case we define $S_n(k) = ||X_n||$. Also, we define $S_n(k) = 0$ for $k < 0$. Note that, in the transient case, the shifts count L_n is precisely equal to the extent of X_n, i.e., $L_n = |X_n|$; indeed, $X_0 = [1]$, so $|X_0| = 0 = L_0$, and, for $n > 0$, $|X_n| - |X_{n-1}| = \mathbf{1}_{B_{n-1}} = L_n - L_{n-1}$. It is convenient to change the indexing of the components of X_n (the bins) from "backwards" to "forwards", so we let

$$\hat{X}_n(j) = X_n(-(L_n - j)).$$

In other words, the configuration is now denoted $\hat{X}_n = [\hat{X}_n(0), \ldots, \hat{X}_n(L_n)]$, for the transient infinite bin model, and $\hat{X}_n = [\ldots, \hat{X}(-1), \hat{X}_n(0), \ldots, \hat{X}_n(L_n)]$, for the stationary infinite bin model. Correspondingly, we let $\hat{S}_n(j) = \hat{X}_n(0) + \cdots + \hat{X}_n(j)$. Note that, in the above, the index j ranges between 0 and L_n in the transient case, and over $\mathbf{Z}_+$ in the stationary case. In the transient case, $\hat{S}_n(L_n) = S_n(L_n) = ||X_n|| = n + 1$. Finally,

$$\hat{S}_n(J) = \sum_{j \in J} \hat{X}_n(j),$$

for any set J of integers, with the understanding $\hat{S}_n(J) = 0$ for $J \subseteq \{-1, -2, \ldots\}$. We can think of the limit theorems in this and the next section, as limit theorems for space-time rescalings of the random measure on the integers defined by this $\hat{S}_n$.

Next, consider the shift counts L_n evaluated at a decoupling epoch $n = \gamma_r$. We obtain the random indices

$$\cdots < L_{\gamma_{-1}} < L_{\gamma_0} \le 0 < L_{\gamma_1} < L_{\gamma_2} < \cdots ,$$

where the indexing follows from that of the previous indexing conventions. As usually, only the doubly-infinite sequence L_{γ_r} is considered for the stationary case, whereas, for the transient case, we only consider L_{γ_r} for $r > 0$. We call these indices *special bins*. Clearly,

$$\frac{L_{\gamma_r}}{r} \to C\rho^{-1}, \text{ a.s.},$$

as $r \to \infty$, for the transient case, or as $r \to \pm\infty$, for the stationary case. The main reason for introducing the special bins is because of the following simple but important computation.

Lemma 8.1. *Let α be a decoupling epoch. Then for all n,*

$$\hat{S}_n[L_\alpha, L_n] = (n - \alpha + 1)^+,$$

both for the transient and stationary case.

Proof. If $n < \alpha$, then $L_n < L_\alpha$, strict inequality, because there is a shift at time α. Hence the interval $[L_\alpha, L_n]$ is empty. Now suppose $n \geq \alpha$. Then

$$\hat{S}_n[L_\alpha, L_n] = \sum_{L_\alpha \leq j \leq L_n} \hat{X}_n(j) = \sum_{0 \leq k \leq L_n - L_\alpha} X_n(-k).$$

Notice, here, that $L_n - L_\alpha = L(\alpha, n]$. The key here is the renovation condition (7.8), with $\ell = 0$, from which it follows

$$X_n = [X_{\alpha-1}, Y_{n-\alpha+1} \circ \theta^{\alpha-1}], \qquad \text{a.s. on } \{n \geq \alpha\}, \tag{8.3}$$

where (recall from (7.5)) Y is defined through

$$Y_1 = [1], \quad Y_{n+1} = \Phi(Y_n, \xi_{n+1}), \quad n \geq 1. \tag{8.4}$$

So Y itself is the transient infinite bin model (but starting at $n = 1$). Hence $||Y_n|| = n$, and $|Y_n| = L_n - L_1 = L(1, n]$. So, for $n \geq \alpha$, $|Y_{n-\alpha+1}| = L(1, n - \alpha + 1]$, and hence

$$|Y_{n-\alpha+1} \circ \theta^{\alpha-1}| = L(1, n - \alpha + 1] \circ \theta^{\alpha-1} = L(\alpha, n]. \tag{8.5}$$

Looking at the last expression for $\hat{S}_n[L_\alpha, L_n]$ we see that it involves $X_n(-k)$ for k up to $L(\alpha, n]$ which is the extent of $Y_{n-\alpha+1} \circ \theta^{\alpha-1}$, and which, by (8.3), is itself the leftmost part of the configuration X_n. Hence

$$\hat{S}_n[L_\alpha, L_\beta) = \sum_{0 \leq k \leq |Y_{n-\alpha+1} \circ \theta^{\alpha-1}|} Y_{n-\alpha+1} \circ \theta^{\alpha-1}(-k) = ||Y_{n-\alpha+1}|| = n - \alpha + 1.$$

□

Let us take another look at $\hat{X}_n$, the configuration of the infinite bin model, in the new indexing. According to this indexing, we can think of bins as having labels that do not change with n. So, a bin that has label k at time n, will continue having label k for ever. (This is in contrast to the previous indexing, according to which a bin changes label every time there is a shift.) In particular, a special bin L_{γ_r} first gets filled with a particle at the decoupling epoch γ_r. Given any bin label $k \in \mathbf{Z}$, there is a special bin with maximal label $\leq k$ and a special bin with minimal label $> k$. The special bin with maximal label $\leq k$ is first filled with a new particle at a decoupling epoch denoted by $M(k)$. This is given by

$$M(k) = \sup\{n \in \mathbf{Z} : \ L_{\gamma(\Gamma_n)} \leq k\}.$$

Note that we use γ_r and $\gamma(r)$, interchangeably, to avoid stacking of indices. The verbal description preceding the last display is only meant to give a physical meaning to this definition. Mathematically though, $M(k)$ *is* defined by the last display, and in fact it is defined for all $k \in \mathbf{Z}$, if we are talking about the stationary model. Let us make all that precise by proving

Lemma 8.2. *The process $\{M(k), k \in \mathbf{Z}\}$ is non-decreasing, tends to $\pm\infty$ as $k \to \pm\infty$, respectively, and, for all $k \in \mathbf{Z}$, $M(k)$ is a decoupling epoch such that*

$$M(k) = \gamma(\Gamma_{M(k)}), \quad L_{\gamma(\Gamma_{M(k)})} \le k < L_{\gamma(1+\Gamma_{M(k)})}.$$

In addition, for all n,

$$M(L_n) = \gamma(\Gamma_n),$$

and $M(k)/k \to C^{-1}$ as $k \to \pm\infty$.

Proof. It is clear that if $k < \ell$ then $M(k) \le M(\ell)$. The supremum in the definition of $M(k)$ is obviously a maximum. Hence $L_{\gamma(\Gamma_{M(k)})} \le k$. The set $\{\gamma(\Gamma_n), n \in \mathbf{Z}\}$ is the same as $D = \{\gamma_r, r \in \mathbf{Z}\}$, i.e., all the decoupling epochs, just because the range of $n \mapsto \Gamma_n$ is $\mathbf{Z}$. Hence the supremum in the definition of $M(k)$ is achieved by an element of D, i.e., $M(k) \in D$. Now, Γ_n is also given by $\Gamma_n = \max\{r : \ \gamma_r \le n\}$. So $\Gamma_{\gamma_r} = r$, and $\gamma(\Gamma_{\gamma_r}) = \gamma_r$. In other words, $\gamma(\Gamma_\alpha) = \alpha$, for all $\alpha \in D$. But $M(k)$ itself is an element of D, so $\gamma(\Gamma_{M(k)}) = M(k)$. The inequality $k < L_{\gamma(1+\Gamma_{M(k)})}$ follows from the fact that L_{γ_r} is strictly smaller than $L_{\gamma_{r+1}}$ for all r. To prove the last equality, note that, for all n, $\gamma(\Gamma_n) \le n < \gamma(1+\Gamma_n)$, and so, $\max\{r : L_{\gamma_r} \le L_n\} = \Gamma_n$. Finally, the limit of $M(k)/k$ is obtained from the fact that $\gamma(\Gamma_n)/n \to 1$, because $r \mapsto \gamma_r$ and $n \mapsto \Gamma_n$ are (generalized) inverses of one another, and the fact that $k \mapsto M(k)$ is (generalized) inverse of $n \mapsto L_n$; the latter satisfies $L_n/n \to C$. □

Let us now take a close look at the point processes of decoupling epochs $\{\gamma_r\}$ and special bins $\{L_{\gamma_r}\}$. They are jointly stationary ergodic on our $(\Omega, \mathcal{F}, \mathsf{P})$, in the sense that $\sum_{r\in\mathrm{Z}} \mathbf{1}(\gamma_r \in J) \circ \theta^n = \sum_{r\in\mathrm{Z}} \mathbf{1}(\gamma_r \in J+n)$, $\sum_{r\in\mathrm{Z}} \mathbf{1}(L_{\gamma_r} \in J) \circ \theta^n = \sum_{r\in\mathrm{Z}} \mathbf{1}(L_{\gamma_r} \in J+n)$, for all $n \in \mathbf{Z}$. Denote by P^0 the conditional probability measure

$$\mathsf{P}^0 := \mathsf{P}(\cdot \mid A_0),$$

where, as usual, $A_0 = \{\xi_1 = 1, \xi_2 \le 2, \xi_3 \le 3, \ldots\}$. Then standard ergodic theory shows that $\vartheta := \theta^{\gamma_1}$ is ergodic and preserves P^0, and the bivariate process $\{(\gamma_{r+1} - \gamma_r, L_{\gamma_{r+1}} - L_{\gamma_r}), r \in \mathbf{Z}\}$ is stationary-ergodic, in the sense that $(\gamma_{r+1} - \gamma_r, L_{\gamma_{r+1}} - L_{\gamma_r}) \circ \vartheta^s = (\gamma_{r+s+1} - \gamma_{r+s}, L_{\gamma_{r+s+1}} - L_{\gamma_{r+s}})$, for all $r, s \in \mathbf{Z}$. We collect these observations, together with some estimates in the next lemma. We can think of this as referring to the stationary infinite bin model. Obvious modifications on the range of indices make it valid for the transient model as well.

Lemma 8.3. *Under P^0, the bivariate process*

$$\{(\gamma_{r+1} - \gamma_r, L_{\gamma_{r+1}} - L_{\gamma_r}), r \in \mathbf{Z}\}$$

is stationary-ergodic, and, if E^0 denotes expectation with respect to P^0,

$$\mathsf{E}^0(\gamma_{r+1} - \gamma_r) = \rho^{-1}, \quad \mathsf{E}^0(L_{\gamma_{r+1}} - L_{\gamma_r}) = C\rho^{-1}.$$

For any positive integers a, b it holds that

$$\max_{-an \le r \le bn} (\gamma_{r+1} - \gamma_r) = o(n), \qquad \max_{-an \le r \le bn} (L_{\gamma_{r+1}} - L_{\gamma_r}) = o(n),$$
$$\max_{-an \le r \le bn} |L_{\gamma_r} - C\gamma_r| = o(n), \qquad \max_{-an \le m \le bn} |L_m - Cm| = o(n),$$

as $n \to \infty$, both P^0 and P-a.s.

Proof. The stationarity and ergodicity statements are standard results. The first two maximal estimates follow from Proposition A.1 of the Appendix. The last estimate follows from Proposition A.2 of the Appendix. Indeed, fix $r > 0$ and write $L_{\gamma_r} - C\gamma_r = \sum_{s=1}^r \zeta_s$, P^0-a.s., where $\zeta_s := (L_{\gamma_s} - L_{\gamma_{s-1}}) - C(\gamma_s - \gamma_{s-1})$, and where we also use the fact that $\mathsf{P}^0(\gamma_0 = 0) = 1$. Now, $\{\zeta_s\}$ is a stationary-ergodic process on $(\Omega, \mathcal{F}, \mathsf{P}^0)$ with $\mathsf{E}^0(\zeta_s) = 0$, and this makes Proposition A.2 applicable. It gives that $\max_{0 \le r \le bn} |L_{\gamma_r} - C\gamma_r| = o(n)$, P^0-a.s. It is easy to see that this holds P-a.s. also. Similarly, we argue for negative indices. For the last estimate, we write

$$\begin{aligned} |L_m - Cm| &\le |L_{\gamma_{\Gamma_m}} - C\gamma_{\Gamma_m}| + |L_m - L_{\gamma_{\Gamma_m}}| + C|m - \gamma_{\Gamma_m}| \\ &\le |L_{\gamma_{\Gamma_m}} - C\gamma_{\Gamma_m}| + |L_{\gamma_{1+\Gamma_m}} - L_{\gamma_{\Gamma_m}}| + C|\gamma_{1+\Gamma_m} - \gamma_{\Gamma_m}|, \end{aligned}$$

where we used $\gamma_{\Gamma_m} \le m < \gamma_{1+\Gamma_m}$, $L_{\gamma_{\Gamma_m}} \le L_m < L_{\gamma_{1+\Gamma_m}}$, for all m. Thus,

$$\max_{-an \le m \le bn} |L_m - Cm| \le \max |L_{\gamma_r} - C\gamma_r| + \max |L_{\gamma_{r+1}} - L_{\gamma_r}| + \max |\gamma_{r+1} - \gamma_r|,$$

where the three maxima on the left are taken over $\Gamma_{-an} \le r \le \Gamma_{bn}$. Use now the previous estimate together with the fact that $\Gamma_{-an}/n \to -a\rho$, $\Gamma_{bn}/n \to b\rho$, as $n \to \infty$, P and P^0-a.s., to conclude. □

Proof of Theorem 8.1 for the stationary infinite bin model

We will show that $\max_{0 \le t \le T} |S_n([nCt]) - nt|$ is $o(n)$, a.s., for any $T > 0$. Now, this quantity is bounded from above by a constant plus

$$\max_{0 \le k \le bn} |S_n(k) - C^{-1}k|,$$

where $b = [CT] + 1$. So we look at the last quantity and show that, for any $b > 0$, it is $o(n)$, a.s. In fact, without loss of generality, we are going to assume that $b > C$. Let

$$\alpha \equiv \alpha_n = M(L_n - bn),$$

so, $L_\alpha \le L_n - bn \le L_n$. Since $b > C$, we have $L_n - bn \to -\infty$, as $n \to \infty$, a.s., and

$$\frac{\alpha}{n} = \frac{M(L_n - bn)}{L_n - bn} \frac{L_n - bn}{n} \to C^{-1}(C - b) = 1 - C^{-1}b < 0,$$

as $n \to \infty$, a.s. So α has a linear growth rate and converges to $-\infty$. We write, for $0 \le k \le bn$,

$$\begin{aligned} S_n(k) &= X_n(-k) + \cdots + X_n(0) \\ &= \hat{X}_n(L_n - k) + \cdots + \hat{X}_n(L_n) \\ &= [\hat{X}_n(L_\alpha) + \cdots + \hat{X}(L_n)] - [\hat{X}_n(L_\alpha) + \cdots + \hat{X}(L_n - k - 1)] \\ &= \hat{S}_n[L_\alpha, L_n] - \hat{S}_n[L_\alpha, L_n - k). \end{aligned}$$

Correspondingly, we write

$$k = (L_n - L_\alpha) - (L_n - L_\alpha - k),$$

and so we have

$$\begin{aligned} |S_n(k) - C^{-1}k| &\le |\hat{S}_n[L_\alpha, L_n] - C^{-1}(L_n - L_\alpha)| \\ &\quad + |\hat{S}_n[L_\alpha, L_n - k) - C^{-1}(L_n - L_\alpha - k)| \\ &=: \mathfrak{A}_n + \mathfrak{B}_{n,k}. \end{aligned}$$

We show that the first term is $o(n)$ and the second is $o(n)$ uniformly in $k \in [0, bn]$. The first term: we have $L_\alpha \le L_n$. Since α is a point of increase of L, it follows that $\alpha \le n$. Hence, from Lemma 8.1 we have $\hat{S}_n[L_\alpha, L_n] = n - \alpha + 1$, and so,

$$C\,\mathfrak{A}_n = |(L_n - L_\alpha) - C(n - \alpha + 1)| \le |L_n - Cn| + |L_\alpha - C\alpha| + C.$$

The terms on the right are $o(n)$, by Lemma 8.3 and the fact that $\alpha/n \to -(1 - C^{-1}b)$. The second term:

$$\begin{aligned} \max_{0 \le k \le bn} \mathfrak{B}_{n,k} &= \max_{L_n - bn \le k \le L_n} |\hat{S}_n[L_\alpha, k) - C^{-1}(k - L_\alpha)| \\ &\le \max_{L_\alpha \le k \le L_n} |\hat{S}_n[L_\alpha, k) - C^{-1}(k - L_\alpha)| \\ &\le \max_{L_\alpha \le k \le L_n} |\hat{S}_n[L_\alpha, k) - (M(k) - \alpha)| \\ &\quad + \max_{L_\alpha \le k \le L_n} |(M(k) - \alpha) - C^{-1}(k - L_\alpha)| \\ &=: \mathfrak{C}_n + \mathfrak{D}_n. \end{aligned}$$

Since $L_\alpha \le k \le L_n$, we have $\alpha \le M(k) \le n$. So $L_\alpha \le L_{M(k)}$ and, by Lemma 8.2, $L_{M(k)} \le k$. Write then

$$\hat{S}_n[L_\alpha, k) = \hat{S}_n[L_\alpha, L_{M(k)}) + \hat{S}_n[L_{M(k)}, k).$$

But $\hat{S}_n[L_\alpha, L_{M(k)}) = \hat{S}_n[L_\alpha, L_{M(k)}] - 1 = M(k) - \alpha$, as we showed in Lemma 8.1. And, since $\hat{S}_n[L_{M(k)}, k) \ge 0$, we get

$$\hat{S}_n[L_\alpha, k) - (M(k) - \alpha) \ge 0.$$

On the other hand,

$$\hat{S}_n[L_\alpha, k) - (M(k) - \alpha) = \hat{S}_n[L_{M(k)}, k) \le \hat{S}_n[L_{\gamma(\Gamma_{M(k)})}, L_{\gamma(1+\Gamma_{M(k)})}) \\ \le \gamma(1 + \Gamma_{M(k)}) - \gamma(\Gamma_{M(k)}).$$

Hence

$$\mathfrak{C}_n \le \max_{\alpha \le M(k) \le n} (\gamma(1 + \Gamma_{M(k)}) - \gamma(\Gamma_{M(k)})) \le \max_{\Gamma_\alpha \le r \le \Gamma_n} (\gamma_{r+1} - \gamma_r),$$

and this is $o(n)$, by Lemma 8.3 and the convergences $\Gamma_n/n \to \rho > 0$, $\Gamma_\alpha/n \to -\rho(1 - C^{-1}b) < 0$. Let us finally look at the term $\mathfrak{D}_n$. We split it again in two parts

$$\begin{aligned} C\mathfrak{D}_n &\le \max_{L_\alpha \le k \le L_n} |(L_{M(k)} - L_\alpha) - C(M(k) - \alpha)| \\ &\quad + \max_{L_\alpha \le k \le L_n} |(L_{M(k)} - L_\alpha) - (k - L_\alpha)| \\ &=: \mathfrak{E}_n + \mathfrak{F}_n. \end{aligned}$$

For the first term we have

$$\mathfrak{E}_n \le \max_{\alpha \le j \le n} |(L_j - L_\alpha) - C(j - \alpha)|,$$

and this is again $o(n)$, by Lemma 8.3. For the other term we have

$$\begin{aligned} \mathfrak{F}_n = \max_{L_\alpha \le k \le L_n} (k - L_{M(k)}) &\le \max_{L_\alpha \le k \le L_n} (L_{\gamma(1+\Gamma_{M(k)})} - L_{\gamma(\Gamma_{M(k)})}) \\ &\le \max_{\Gamma_\alpha \le r \le \Gamma_n} (L_{\gamma_{r+1}} - L_{\gamma_r}), \end{aligned}$$

which is $o(n)$, again by Lemma 8.3. □

Proof of Theorem 8.1 for the transient infinite bin model

We continue with the transient bin model. To differentiate between the transient and stationary case, we will use tildes for the stationary one. So far we proved

$$\max_{0 \le k \le bn} |\tilde{S}_n(k) - C^{-1}k| = o(n),$$

for any $b > 0$. Consider the first decoupling epoch $\gamma = \gamma_1$. Let $\tilde{L}(\gamma, n]$ be the number of shifts on the interval $(\gamma, n]$, for the stationary model. Since $\tilde{L}(\gamma, n]/n \to C$, we also have that

$$\tilde{U}_n := \max_{0 \le k \le \tilde{L}(\gamma, n]} |\tilde{S}_n(k) - C^{-1}k| = o(n).$$

Consider (8.3), and write it, using tildes, as

$$\tilde{X}_n = [\tilde{X}_{\gamma-1}, Y_{n-\gamma+1}\circ\theta^{\gamma-1}], \qquad \text{a.s. on } \{n \ge \gamma\}.$$

Recall that Y satisfies (8.4). So if we let $X_n = Y_{n+1}\circ\theta^{-1}$, we see that X satisfies

$$X_0 = [1], \quad X_{n+1} = \Phi(X_n, \xi_{n+1}), \quad n \ge 0.$$

Hence X is the transient infinite bin model. We then write

$$\tilde{X}_n = [\tilde{X}_{\gamma-1}, X_{n-\gamma}\circ\theta^{\gamma}], \qquad \text{a.s. on } \{n \ge \gamma\}.$$

From (8.5) we get $|X_{n-\gamma}\circ\theta^{\gamma}| = \tilde{L}(\gamma, n]$. But $|X_n| = L_n$ (the number of shifts is the extent in the transient model). So $\tilde{L}(\gamma, n] = L_{n-\gamma}\circ\theta^{\gamma}$. Thus,

$$\begin{aligned}\tilde{U}_n &= \max_{0\le k\le \tilde{L}(\gamma,n]} |\tilde{S}_n(k) - C^{-1}k| = \max_{0\le k\le \tilde{L}(\gamma,n]} \left| \sum_{j=0}^{k} (X_{n-\gamma}\circ\theta^{\gamma})(-j) - C^{-1}k \right| \\ &= \max_{0\le k\le L_{n-\gamma}\circ\theta^{\gamma}} |S_{n-\gamma}(k)\circ\theta^{\gamma} - C^{-1}k| = U_{n-\gamma}\circ\theta^{\gamma},\end{aligned}$$

where

$$U_n := \max_{0\le k\le L_n} |S_n(k) - C^{-1}k|.$$

Since $\tilde{U}_n/n \to 0$, P-a.s., we have $U_{n-\gamma}\circ\theta^{\gamma}/n \to 0$, P-a.s. Since $\gamma < \infty$, P-a.s., we have $U_n\circ\theta^{\gamma}/n \to 0$, P-a.s. By ergodicity, $U_n\circ\theta^{\gamma}/n \to 0$, P^0-a.s. Since θ^{γ} preserves P^0, we have $U_n/n \to 0$, P^0-a.s. And finally, by ergodicity again, $U_n/n \to 0$, P-a.s., also. It is now easy to translate this into the statement $\max_{0\le t\le T} |S_n([nCt]) - n(t\wedge 1)| = o(n)$. For $T = 1$, we have

$$\begin{aligned}\max_{0\le t\le 1} |S_n([nCt]) - n(t\wedge 1)| &\le \max_{0\le k\le [nC]} |S_n(k) - C^{-1}k| + C^{-1} \\ &= U_n \vee \max_{L_n\le k\le [nC]} |n - C^{-1}k| + C^{-1},\end{aligned}$$

where the last equality follows from a splitting of the maximum and the fact that $S_n(k) = n$ for $k \ge L_n$, for the transient model. The last maximum is bounded above by a constant times $|L_n - nC|$ which is $o(n)$, as shown earlier. Finally, for $T > 1$, we only have to look at

$$\max_{1\le t\le T} |S_n([nCt]) - n(t\wedge 1)| \le \max_{[nC]\le k\le [nCT]} |S_n(k) - C^{-1}[nC]| + C^{-1}.$$

We can then write, on the event $\{L_n \le [nCT]\}$ (the probability of which converges to 1),

$$\begin{aligned}&\max_{[nC]\le k\le [nCT]} |S_n(k) - C^{-1}[nC]| \\ &\qquad \le \max_{[nC]\le k\le L_n} |S_n(k) - C^{-1}[nC]| \vee \max_{L_n\le k\le [nCT]} |S_n(k) - C^{-1}[nC]|.\end{aligned}$$

In the second term, $S_n(k) = n$, so this term is bounded by a constant. The first term is bounded by a constant plus U_n. Hence they are both $o(n)$. □

9. Functional central limit theorem for the infinite bin model

Now we study the deviation of $S_n([nCt])$ from its functional mean, which, according to the results of the previous section, is a linear function. We consider the stationary infinite bin model, and, to make things simple, we assume that the ξ_n are i.i.d.

Lemma 9.1. *Consider the stationary infinite bin model. If the driving sequence $\{\xi_n\}$ is i.i.d., nontrivial, and $\mathsf{E}\xi_0 < \infty$, then the point process of decoupling epochs $\{\gamma_r\}$ is a stationary renewal process. Also, the point process of special bins $\{L_{\gamma_r}\}$ is another stationary renewal process.*

Proof. From the theory of stationary point processes, it suffices to show that, conditionally on $\{\gamma_0 = 0\}$, the sequence $\{\gamma_{r+1} - \gamma_r,\ r \in \mathbf{Z}\}$ is i.i.d. In fact, more is true: conditionally on $\{\gamma_0 = 0\}$ the bivariate sequence $\{(\gamma_{r+1} - \gamma_r, L_{\gamma_{r+1}} - L_{\gamma_r}),\ r \in \mathbf{Z}\}$ is i.i.d. From the definition of the sequence $\{\gamma_r\}$, we have

$$\begin{aligned}\gamma_{-1} &= \sup\{n < \gamma_0 :\ \mathbf{1}_{A_{n-1}} = 1\} \\ &= \sup\{n < \gamma_0 :\ \xi_n = 1, \xi_{n+1} \le 2, \ldots, \xi_{-1} \le -n, \\ &\qquad \xi_0 \le -n+1, \xi_1 \le -n+2, \ldots\}.\end{aligned}$$

Consider the event $\{\gamma_0 = 0\}$. We have $\{\gamma_0 = 0\} = A_{-1} = \{\xi_0 \le 1, \xi_1 \le 2, \xi_2 \le 3, \ldots\}$. So, conditionally on this event, γ_{-1} is distributed as

$$\sup\{n < 0 :\ \xi_n = 1, \xi_{n+1} \le 2, \ldots, \xi_{-1} \le -n\},$$

and is independent of $\{\gamma_r, r \ge 1\}$. Moreover, conditionally on $\{\gamma_0 = 0\}$, the array $\{\xi_n\}_{\gamma_{-1} \le n < 0}$ is independent of $\{\xi_n, n \ge 0\}$. Since $L_{\gamma_{r+1}} - L_{\gamma_r}$ is a deterministic function of $\{\xi_n\}_{\gamma_r \le n < \gamma_{r+1}}$, for all r, the result follows. $\square$

Note that the distribution of $\gamma_{r+1} - \gamma_r$, conditionally on $\{\gamma_0 = 0\}$, is the distribution of the stopping time

$$\inf\big\{n \ge 1 :\ \max_{1 \le k \le n} (\xi_k + k - 1) \le n\big\}, \tag{9.1}$$

which is, in principle, computable. But what is of interest is the distribution of $L_{\gamma_{r+1}} - L_{\gamma_r}$. This is a harder problem. More on that in the next section.

Lemma 9.2. *Under the assumptions above, if $\{r(n), n = 1, 2, \ldots\}$ is a random sequence such that $r(n)/n \to c > 0$, a.s., then the random sequence*

$$(\gamma_{r(n)+1} - \gamma_{r(n)},\ L_{\gamma_{r(n)+1}} - L_{\gamma_{r(n)}}), \quad n = 1, 2, \ldots$$

is tight.

Proof. Omitted.

Theorem 9.1. *Suppose that the driving sequence $\{\xi_n\}$, for the infinite bin model, is i.i.d., with values in* $\mathbf{N}$*, such that* $\mathsf{P}(\xi_0 = 1) < 1$ *and finite mean. Let* $\{X_n, n \in \mathbf{Z}\}$ *be the stationary version of the infinite bin model, and* $S_n(k) = X_n(-k) + \cdots + X_n(0), k \geq 0$*. Define*

$$\eta_n(t) = n^{-1/2}\big(S_n([nCt]) - nt\big), \quad t \geq 0,$$

and consider $\{\eta_n, n = 1, 2, \ldots\}$ *as a sequence of random elements of* $D[0,\infty)$ *with the uniform topology on compacta. As such, it converges weakly to the process* $\{\sqrt{\rho}\sigma W_t, t \geq 0\}$*, where* ρ *is the renovation rate (8.1),* $\sigma^2 = \mathrm{var}[\gamma_1 - \gamma_0 - C^{-1}(L_{\gamma_1} - L_{\gamma_0})]$*, and* W *is a standard Brownian motion.*

Proof. Define

$$K \equiv K_{n,t} = [nCt], \quad \alpha \equiv \alpha_{n,t} = M(L_n - [nCt]).$$

Wherever there is no confusion, we omit the dependence on t or n. It suffices to show that $\{n^{-1/2}(S_n(K_{n,t}) - C^{-1}K_{n,t}),\ t \geq 0\}$ converges to a Brownian motion. We have

$$S_n(K) = X_n(-K) + \cdots + X_n(0) = \hat{X}_n(-K) + \cdots + \hat{X}_n(0) = \hat{S}_n[L_n - K, L_n].$$

Decompose this as follows

$$\hat{S}_n[L_n - K, L_n] = \hat{S}_n[L_\alpha, L_{\gamma(\Gamma_n)}) - \hat{S}_n[L_\alpha, L_n - K) + \hat{S}_n[L_{\gamma(\Gamma_n)}, L_n].$$

Correspondingly, decompose K as

$$K = (L_{\gamma(\Gamma_n)} - L_\alpha) - (L_n - K - L_\alpha) + (L_n - L_{\gamma(\Gamma_n)}).$$

Hence $S_n(K) - C^{-1}K$ is the sum of three terms

$$\begin{aligned} \mathfrak{z}_n(t) &:= \hat{S}_n[L_\alpha, L_{\gamma(\Gamma_n)}) - C^{-1}(L_{\gamma(\Gamma_n)} - L_\alpha), \\ \mathfrak{z}'_n(t) &:= \hat{S}_n[L_\alpha, L_n - K) - C^{-1}(L_n - K - L_\alpha), \\ \mathfrak{z}''_n &:= \hat{S}_n[L_{\gamma(\Gamma_n)}, L_n] - C^{-1}(L_n - L_{\gamma(\Gamma_n)}). \end{aligned}$$

Consider the third term. Use the inequalities $n < \gamma(1+\Gamma_n)$, and $L_n \leq L_{\gamma(1+\Gamma_n)}$, to write

$$|\mathfrak{z}''_n| \leq (\gamma(1+\Gamma_n) - \gamma(\Gamma_n)) + C^{-1}(L_{\gamma(1+\Gamma_n)} - L_{\gamma(\Gamma_n))}).$$

Both terms are tight random sequences. Hence $n^{-1/2}\mathfrak{z}''_n$ converges to zero in probability. Similarly, we show that $\mathfrak{z}''_n$ converges to zero in probability. Use an inequality from Lemma 8.2

$$L_n - K < L_{\gamma(1+\Gamma_\alpha)},$$

the fact that $\alpha = \gamma(\Gamma_\alpha)$, to get

$$\hat{S}_n[L_\alpha, L_n - K] \le \hat{S}_n[L_{\gamma(\Gamma_\alpha)}, L_{\gamma(1+\Gamma_\alpha)}) = \gamma(1+\Gamma_\alpha) - \gamma(\Gamma_\alpha),$$

where the last equality follows from Lemma 8.1. Hence we get

$$|\mathfrak{z}'_n(t)| \le \big(\gamma(1+\Gamma_\alpha) - \gamma(\Gamma_\alpha)\big) + C^{-1}\big(L_{\gamma(1+\Gamma_\alpha)} - L_{\gamma(\Gamma_\alpha)}\big),$$

and both terms are again tight. Let us write the first term, $\mathfrak{z}_n(t)$, as follows

$$\mathfrak{z}_n(t) = \sum_{r=a+1}^{b} \chi_r,$$

where

$$a \equiv a_{n,t} = \Gamma_\alpha = \Gamma_{M(L_n - [nCt])}, \quad b \equiv b_n = \Gamma_n,$$
$$\chi_r = (\gamma_r - \gamma_{r-1}) - C^{-1}(L_{\gamma_r} - L_{\gamma_{r-1}}).$$

From Lemmas 9.2 and 8.3 the random variables $\{\chi_r\}$ are i.i.d. with zero mean. We have that the random sequence $W_n(t) := n^{-1/2}\sum_1^{[nt]} \xi_r$ (from Donsker's theorem) converges weakly to σW_t, where W is a standard Brownian motion, and $\sigma^2 = \mathrm{var}(\chi_r)$. Define

$$\varphi_n(t) = \frac{\Gamma_n - \Gamma_{M(L_n - [nCt])}}{n}, \quad t \ge 0.$$

Observe that $\varphi_n(0) = 0$, φ_n has paths in $D[0,\infty)$ (in fact, they are increasing), and φ_n converges, uniformly in probability, to the function ρt. On the other hand, $n^{-1/2}\mathfrak{z}_n$ is identical in distribution to $W_n \circ \varphi_n$. By the continuous mapping theorem, we obtain that $n^{-1/2}\mathfrak{z}_n$ converges to $\sqrt{\rho}\sigma W$. □

Theorem 9.2. *Under the same assumptions of Theorem* 9.1, *consider, for* $n = 1, 2, \dots,$

$$\eta_n^0(t) := n^{-1/2}\big(S_n([tL_n]) - nt\big), \quad 0 \le t \le 1.$$

Then, as $n \to \infty$, η_n^0 *converges weakly to a Brownian Bridge* (*BB*).

Before proving the theorem, let us make some remarks. Here, the quantity L_n refers to the number of shifts on the interval $(0, n]$, as defined in Section 8.1. Recall that γ_0 is the first decoupling epoch prior to 0, and γ_1 the first such epoch after 0; we have

$$\gamma_0 \le 0 < \gamma_1, \quad L_{\gamma_0} \le 0 < L_{\gamma_1},$$

as already noted in Section 8.1. Hence,

$$\hat{S}_n[L_{\gamma_0}, L_n] \ge S_n(L_n) = \hat{S}_n[0, L_n] \ge \hat{S}[L_{\gamma_1}, L_n],$$

(for $n \ge \gamma_1$). But $\hat{S}_n[L_{\gamma_0}, L_n] = n - \gamma_0 + 1$, and $\hat{S}[L_{\gamma_1}, n] = n - \gamma_1 + 1$ (see Lemma 8.1). Hence

$$|S_n(L_n) - n| \le 1 + |\gamma_0| \vee \gamma_1,$$

and so

$$\eta_n^0(1) \to 0, \quad \text{as } n \to \infty.$$

So we know, at least, that the values at the end points $t = 0$, and $t = 1$ are zero, in the limit, as they should!

Proof. Decompose $\eta_n^0(t)$ as

$$\eta_n^0(t) = \left[\frac{S_n([tL_n]) - C^{-1}tL_n}{\sqrt{n}}\right] - t\left[\frac{n - C^{-1}L_n}{\sqrt{n}}\right] =: \alpha_n(t) - t\delta_n.$$

Note that if we define

$$\xi_n(t) := \frac{S_n([nCt]) - nt}{\sqrt{n}}, \quad \varphi_n(t) := \frac{tL_n}{nC},$$

we have

$$\alpha_n = \xi_n \circ \varphi_n.$$

But, from Theorem 9.1,

$$\xi_n \Rightarrow \sqrt{\rho}\sigma W,$$

where W is a standard Brownian motion. Also, $\varphi_n(t) \to t$, *uniformly in* $t \in [0, 1]$, from the functional law of large numbers (Theorem 8.1). Since composition is continuous in the Skorokhod topology (see, e.g., [5]), we have that

$$\alpha_n \Rightarrow \sqrt{\rho}\sigma W,$$

also. On the other hand,

$$|\delta_n - \alpha_n(1)| = \left|\frac{S_n(L_n) - n}{\sqrt{n}}\right|,$$

and this converges to zero, as argued in the remarks above. Hence,

$$\delta_n \Rightarrow \sqrt{\rho}\sigma W_1,$$

as $n \to \infty$. Combining the above, we conclude that

$$\{\eta_n^0(t), 0 \le t \le 1\} \Rightarrow \sqrt{\rho}\{W_t - tW_1, 0 \le t \le 1\},$$

which is a Brownian bridge. $\square$

<u>*Final remarks for this section*</u>

1. Analogous theorems hold for the transient case, so we will omit them. We only point out that, due to the nature of L_n in the transient case (it is the extent of the configuration at time n) it may be more natural to consider the scaling of Theorem 9.2 instead that of Theorem 9.1. In other words, it may make more physical sense to approximate the difference of the cumulative sums from their mean behaviour by a Brownian bridge in the transient case.

2. Clearly, the underlying theorems used for the proofs are regenerative central limit theorems. So independence can be dropped and replaced by mixing conditions if desired. The proofs will remain intact.

10. Stochastic ordered graphs

We now consider a stochastic ordered graph with vertex set the integers and directed edges (i, j) with $i < j$ occurring, independently of each other, with probability p. Such a model is of interest in several applications. We mention two. In *Mathematical Ecology*, it models community food webs; see, e.g., [21], and [13]. In *Performance Evaluation* of computer systems, it models task graphs; see, e.g., [18]. Of interest in both these applications are longest paths occurring on the restriction of the stochastic graph on the set $\{1, \ldots, n\}$. Newman [20] derives asymptotics for maximal lengths when the probability p is chosen to depend on n in an inversely proportional fashion.

Our interest is in studying asymptotics for the maximal length of the restriction on $\{1, \ldots, n\}$, as $n \to \infty$, when the connectivity probability p remains constant, and related limit theorems. We do some preliminary work to transform the stochastic ordered graph into an infinite bin model.

Our assumptions are as follows. In the simplest case, we are given a sequence of i.i.d. random variables $\{\alpha_{ij}, -\infty < i < j < \infty\}$, with $\mathsf{P}(\alpha_{i,j} = 1) = p$, $\mathsf{P}(\alpha_{i,j} = -\infty) = 1 - p$. The meaning of $\alpha_{i,j}$ is that it indicates whether there is an edge from i to j (in which case $\alpha_{i,j} = 1$) or not (in which case $\alpha_{i,j} = -\infty$). More generally, let

$$\alpha^n = (\ldots, \alpha_{n-2,n}, \alpha_{n-1,n})$$

and assume that $\{\alpha^n, n \in \mathbf{Z}\}$ are i.i.d. with the property that the marginal distribution, i.e., the distribution of $(\ldots, \alpha_{-3,0}, \alpha_{-2,0}, \alpha_{-1,0})$ is that of an exchangeable random sequence.

The collection of these random variables $\{\alpha_{i,j}\}$ defines the stochastic ordered graph. Consider now a vertex $j > 0$ and define its *weight* W_j to be the length of the longest path that ends at j, and starts at some positive vertex. If j is not connected to any other vertex smaller than j, then $W_j = 0$. In other words, we inductively define $W_1, W_2, \ldots$ by

$$W_1 = 0, \quad W_j = \max_{1 \le i \le j-1} (W_i + \alpha_{i,j})^+, \quad j > 1.$$

We are interested in the random variable

$$L_n = \max(W_1, \ldots, W_n).$$

Let us define

$$X_n(-k) = \sum_{i=1}^{n} \mathbf{1}(W_i = L_n - k), \quad k = 0, 1, \ldots, L_n,$$
$$X_n = [X_n(-L_n), \ldots, X_n(-1), X_n(0)].$$

Thus, $X_n(0)$ is the number of vertices within $\{1, \ldots, n\}$ with maximal weight $(= L_n)$. On the other extreme, $X_n(-L_n)$ is the number of vertices that are not left-connected. Clearly, $X_n(0) + X_n(-1) + \cdots + X_n(-L_n) = n$. We claim that $\{X_n\}$ is identical in distribution to an infinite bin model. Consider a permutation σ_n of $(1, \ldots, n)$ that puts the weights W_i in decreasing order. Since not all weights are necessarily distinct, the permutation is not unique. To fix a particular permutation, let σ_n be the permutation of $(1, \ldots, n)$ that puts the pairs $\{(W_i, i), 1 \le i \le n\}$ in decreasing lexicographic order. In other words,

$$i < j \Rightarrow W_{\sigma_n(i)} > W_{\sigma_n(j)}, \text{ or } \big[W_{\sigma_n(i)} = W_{\sigma_n(j)} \text{ and } \sigma_n(i) > \sigma_n(j)\big].$$

Given the weights $W_1, \ldots, W_n$, the weight W_{n+1} is found by looking at the vertex $i \le n$ such that W_i is largest among $W_1, \ldots, W_n$, and (i, j) is an edge. Since we have ordered the weights, we define

$$\hat{\xi}_{n+1} = \inf\{i \in [1, n] : \ \alpha_{\sigma_n(i), n+1} = 1\}. \tag{10.1}$$

Among the weights $W_{\sigma_n(1)} \ge W_{\sigma_n(2)} \ge \ldots \ge W_{\sigma_n(n)}$, the first $X_n(0)$ of them are equal to L_n, the next $X_n(-1)$ of them are equal to $L_n - 1$, etc. So, if $\hat{\xi}_{n+1} \le X_n(0)$ then $W_{n+1} = \max(W_1, \ldots, W_n) + 1$. Thus, a new maximum is achieved and, in this case, $X_{n+1} = [X_n, 1]$. If $X_n(0) < \hat{\xi}_{n+1} \le X_n(0) + X_n(-1)$, then no new maximum is achieved; rather, the number of vertices with maximal weight increases by 1, and so $X_{n+1} = [X_n(-L_n), \ldots, X_n(-2), X_n(-1), X_n(0) + 1]$. Proceeding this way, we see that X_n satisfies

$$X_{n+1} = \Phi(X_n, \hat{\xi}_{n+1}),$$

where Φ is the same as in (7.2), and $\hat{\xi}_{n+1}$ is defined in (10.1). This is not an infinite bin model, since $\{\xi_n\}$ is not stationary. However, owing to the assumption that α^{n+1} is independent of the past and that the components of α^{n+1} are exchangeable, it follows that $\hat{\xi}_{n+1}$ is identical in distribution to

$$\xi_{n+1} = \inf\{i \in [1, n] : \ \alpha_{i, n+1} = 1\}.$$

The sequence $\{\hat{\xi}_n\}$ is identical in distribution to $\{\xi_n\}$. So $\{X_n\}$ is identical in distribution to the SRS defined by

$$X_{n+1} = \Phi(X_n, \xi_{n+1}).$$

In the simple case, where the $\alpha_{i,j}$ are i.i.d., ξ_n is geometric. Under the exchangeability assumption, more general distributions for ξ_n are possible. By de Finetti's theorem, this distribution must be a mixture of geometric distributions. This is a fairly general class, so we will need to make the assumption that $\mathsf{E}\xi_n < \infty$.

Hence, all the results obtained for the infinite bin model also apply to the stochastic ordered graph model. In particular, we have the existence of renovating events and a unique stationary version for X. We will not try to attribute any special physical meaning to this stationary version. Rather, we shall use it in order to estimate C.

The perfect simulation approach

This approach is based on the perfect simulation algorithm, described in Section 4. We assume that $\mathsf{E}\xi_n < \infty$, so that the renovating events

$$A_n = \{\xi_{n+1} = 1, \xi_{n+2} \le 2, \dots\}$$

have positive probability. Their structure, being of the simple form as required by the perfect simulation algorithm, leads to the definition of the time

$$\beta(0) = \min\{n \ge 0 : \ \xi_{-n} = 1, \xi_{-n+1} \le 2, \dots, \xi_0 \le n+1\}.$$

This is a.s. finite. We can then obtain samples from the stationary $X_0(0)$ by starting at time $-\beta(0)$ from an arbitrary configuration, say the trivial one, and then monitoring X_n for $-\beta(0) < n \le 0$, till we obtain $X_0(0)$. Specifically, we first wait till $\beta(0)$ gets realized. Then we set $X_{-\beta(0)} = [1]$, and, recursively, $X_{n+1} = \Phi(X_n, \xi_{n+1})$, for $-\beta(0) \le n \le 0$. When $n = 0$, we compute $X_0(0)$. This is one (perfect) sample, say S_1, from the stationary distribution. We repeat the same process, independently, obtaining, say, S_j at the jth trial, $j = 1, \dots, m$. We use the formula $1 - \big(\sum_{j=1}^m q^{S_j}\big)/m$ to estimate $C(q)$.

There is a different approach, which may be more advantageous in certain cases. Namely, since we can simulate, not only $X_0(0)$, but any window $[X_0(0), X_1(0), \dots, X_n(0)]$, we may estimate $C(q)$ by $1 - \big(\sum_{j=0}^n q^{X_n(0)}\big)/n$. The way to simulate joint distributions is described in the perfect simulation algorithm of Section 4. However, it may be computationally costly. In addition, combination of the above two approaches is also possible: a way to reduce the computational complexity of the second approach is to take a fixed window of "moderate" size n, and then obtain a number of m independent copies of it. Of course, finding the "optimal" n and m, as well as estimating computational complexity and confidence intervals are problems of interest, but are beyond the scope of this paper.

The advantage of the simulation approach for estimating $C(q)$ is that it can be used also in cases where there are dependencies between the random variables $(\alpha_{i,j},\ i = j-1, j-2, \ldots)$

If the $\alpha_{i,j}$ are i.i.d. with $\mathsf{P}(\alpha_{i,j} = 1) = p$, $\mathsf{P}(\alpha_{i,j} = -\infty) = 1 - p = q$, we develop analytical methods for obtaining sharp upper and lower bounds on $C(q)$. This is done next.

The analytical approach

Consider the stationary version, still denoted by X, for simplicity of notation. The independence assumption implies that X is a Markov process with values in $\mathbf{N}^\infty$. Assume, for the remainder of this section, that the $\alpha_{i,j}$ are i.i.d. with $\mathsf{P}(\alpha_{i,j} = 1) = p$, $\mathsf{P}(\alpha_{i,j} = -\infty) = 1 - p = q$. Hence ξ_0 is geometric with parameter q, i.e.,

$$\mathsf{P}(\xi_0 > k) = q^k, \quad k = 1, 2, \ldots$$

Let $C = \lim_{n\to\infty} L_n/n$. We will prove the following bounds.

Theorem 10.1. *Define the functions*

$$\begin{aligned}
f(q) &:= \sum_{k\geq 1} p^{k-1} q^{(k+1)(k+2)/2} = q^3 + pq^6 + p^2 q^{10} + p^3 q^{15} + \cdots, \\
g(q) &:= \sum_{k\geq 1} q^{k(k+3)/2} = q^2 + q^5 + q^9 + q^{14} + \cdots, \\
L(q) &:= (1-q^2) \max\left\{ \frac{1+q}{1+2q},\ \frac{1+q-q^2-q^4+q^5}{1+2q-q^2-q^4+q^5-q^3+q^6} \right\}, \\
U(q) &:= \min\left\{ \frac{1-pg(q)}{1+q+pf(q)-pg(q)},\ (1-q)\sum_{n=0}^{\infty} \frac{q^{2^n-1}}{1+q^{2^n}} \right\}.
\end{aligned}$$

Then, for $0 < q < 1$,

$$L(q) < C < U(q).$$

To prove this, we need two auxiliary lemmas, which are proved first.

Lemma 10.1. *Consider the events*

$$\begin{aligned}
B_n &= \{\xi_{n+1} \leq X_n(0)\}, \quad n \in \mathbf{Z}, \\
D_n &= \{X_n(0) = 1\}, \quad n \in \mathbf{Z}, \\
H_{-m} &= B_{-m} B^c_{-m+1} \cdots B^c_{-1} D_0, \quad m \geq 2.
\end{aligned}$$

We then have

$$B_n^c D_{n+1} = D_n \cap \{\xi_{n+1} > 1 + X_n(-1)\}, \quad n \in \mathbf{Z}, \tag{10.2}$$

$$H_{-m} \subseteq \{\xi_{-m+1} \le X_{-m}(0)\} \cap \bigcap_{r=1}^{m-1} \{\xi_{-m+r+1} > X_{-m}(0) + 1\}, \quad m \ge 2, \tag{10.3}$$

$$H_{-m} \supseteq \{\xi_{-m+1} \le X_{-m}(0)\} \cap \bigcap_{r=1}^{m-1} \{\xi_{-m+r+1} > X_{-m}(0) + r\}, \quad m \ge 2. \tag{10.4}$$

Proof. Let us first show (10.2). The left-hand side is the event $\{\xi_{n+1} > X_n(0), X_{n+1}(0) = 1\}$. From the definition of Φ, if $\xi_{n+1} > X_n(0)$, then $X_{n+1}(0) = X_n(0) + \mathbf{1}(X_n(0) < \xi_{n+1} \le X_n(0) + X_n(-1))$. If $X_{n+1}(0) = 1$, then the last indicator must be zero, i.e., $\xi_{n+1} > X_n(0) + X_n(-1)$, and so $X_{n+1}(0) = X_n(0)$; so $X_n(0) = 1$. Hence, $B_n^c D_{n+1} \subseteq \{\xi_{n+1} > 1 + X_n(-1)\} \cap \{X_n(0) = 1\}$, which is the right-hand side of (10.2). Conversely, if $X_n(0) = 1$ and $\xi_{n+1} > 1 + X_n(-1)$, then $\xi_{n+1} > X_n(0) + X_n(-1)$, and so, again from the definition of Φ, $X_{n+1}(0) = X_n(0) = 1$. This shows that the opposite inclusion also holds, and so (10.2) has been proved. Consider now $H_{-m} = B_{-m} B_{-m+1}^c \cdots B_{-1}^c D_0$, and apply (10.2), inductively, to get

$$H_{-m} = \{\xi_{-m+1} \le X_{-m}(0)\} \cap \bigcap_{r=1}^{m-1} \{\xi_{-m+r+1} > X_{-m+r}(-1) + 1\}. \tag{10.5}$$

Arguing in a manner similar to the proof of (10.2), we see that, a.s. on H_{-m},

$$\begin{aligned} X_{-m}(0) &= X_{-m+1}(-1) \le X_{-m+2}(-1) \le \cdots \le X_{-1}(-1), \\ X_{-m}(0) &= X_{-m+1}(-1), \\ X_{-m+2}(-1) &\le X_{-m+1}(-1) + 1, \quad \ldots, \\ X_{-1}(-1) &\le X_{-2}(-1) + 1. \end{aligned}$$

Using the first set of inequalities in (10.5) proves inclusion (10.3), and using the second set proves (10.4). □

Lemma 10.2. *For all* q,

$$C(q) \le \frac{1-q}{1+q} + \frac{q}{1+q} C(q^2). \tag{10.6}$$

Proof. We work with the stationary version $\{X_n\}$. From the discussion preceding Theorem 8.1 and, in particular, formula (8.2), we have

$$\begin{aligned} C = C(q) &= \mathsf{P}(\xi_2 \le X_1(0)) \\ &= \mathsf{P}(\xi_2 \le X_1(0), \xi_1 \le X_0(0)) + \mathsf{P}(\xi_2 \le X_1(0), \xi_1 > X_0(0)) \\ &=: P_1 + P_2. \end{aligned}$$

To estimate P_1, observe that if $\xi_1 \le X_0(0)$, then $X_1(0) = 1$, and so

$$P_1 = \mathsf{P}\big(\xi_2 = 1, \xi_1 \le X_0(0)\big) = p\mathsf{E}\big(1 - q^{X_0(0)}\big) = pC(q).$$

As for P_2, if $\xi_1 > X_0(0)$ then $X_0(0) \le X_1(0) \le X_0(0) + 1$. Hence

$$\begin{aligned} P_2 &\le \mathsf{P}\big(\xi_2 \le X_0(0) + 1, \xi_1 > X_0(0)\big) = \mathsf{E}\big[(1 - q^{X_0(0)+1})q^{X_0(0)}\big] \\ &= \mathsf{E}\big[q^{X_0(0)}\big] - q\mathsf{E}\big[q^{2X_0(0)}\big] = \big(1 - C(q)\big) - q\big(1 - C(q^2)\big). \end{aligned}$$

Combining the estimates for P_1 and P_2, the lemma is proved. □

Proof of Theorem 10.1. Let $\{X_n\}$ denote the stationary version of the infinite bin model with i.i.d. geometric driver sequence $\{\xi_n\}$. Denote by

$$\pi_k = \mathsf{P}(X_n(0) = k), \quad k = 1, 2, \dots$$

the marginal distribution of the zeroth component of the steady-state version. From the discussion preceding Theorem 8.1 and, in particular, formula (8.2), we have

$$C = \mathsf{P}(\xi_1 \le X_0(0)) = 1 - \mathsf{E}q^{X_0(0)} = 1 - \sum_{k \ge 1} \pi_k q^k. \tag{10.7}$$

Consider, as earlier, the shifting event

$$B_n = \{\xi_{n+1} \le X_n(0)\},$$

and the event that X_{n+1} has a single particle at its zeroth component

$$D_{n+1} = \{X_{n+1}(0) = 1\}.$$

Clearly,

$$B_n \subseteq D_{n+1},$$

for all n. Now fix n, say $n = 0$, and decompose D_0 as $D_0 = B_{-1}D_0 \cup B^c_{-1}D_0$, which, in view of the inclusion above, becomes

$$D_0 = B_{-1} \cup B^c_{-1}D_0. \tag{10.8}$$

Continuing this, write D_0 as the disjoint union

$$D_0 = B_{-1} \cup \bigcup_{m \ge 2} B_{-m}B^c_{-m+1} \cdots B^c_{-1}D_0. \tag{10.9}$$

As in Lemma 10.1, denote by H_{-m} the generic term on the right,

$$H_{-m} = B_{-m}B^c_{-m+1} \cdots B^c_{-1}D_0, \quad m \ge 2.$$

From (10.8) we get
$$\mathsf{P}(D_0) = \mathsf{P}(B_{-1}) + \mathsf{P}(B_{-1}^c D_0).$$
Recall that $\mathsf{P}(D_0) = \mathsf{P}(X_0(0) = 1) = \pi_1$, $\mathsf{P}(B_{-1}) = C$, and use (10.2) to get
$$\mathsf{P}(B_{-1}^c D_0) \le \mathsf{P}(X_{-1}(0) = 1, \xi_0 > 2) = \pi_1 q^2.$$
Hence,
$$\pi_1 \le C + \pi_1 q^2,$$
which gives an upper bound of π_1 in terms of C:
$$\pi_1 \le \frac{C}{1-q^2}.$$
We already have (see (10.7)) the following trivial bound
$$C \ge 1 - \pi_1 q - q^2 \sum_{k\ge 2} \pi_k = 1 - \pi_1 q - q^2(1-\pi_1) = (1-q^2) - (q-q^2)\pi_1. \quad (10.10)$$
Combining the last two displays we obtain
$$C \ge \frac{(1-q^2)(1+q)}{1+2q} =: L_1(q). \quad (10.11)$$
Now use (10.9) to write
$$\mathsf{P}(D_0) = \mathsf{P}(B_{-1}) + \sum_{m\ge 2} \mathsf{P}(H_{-m}).$$
From (10.3) we also have, upon conditioning on $X_{-m}(0)$ and recalling that $\mathsf{P}(X_{-m}(0) = k) = \pi_k$,
$$\mathsf{P}(H_{-m}) \le \sum_{k\ge 1} \pi_k (1-q^k)(q^{k+1})^{m-1} \le \pi_1(1-q)q^{2(m-1)} + (1-\pi_1)q^{3(m-1)}.$$
Thus,
$$\begin{aligned}\pi_1 &\le C + \sum_{m\ge 2} \mathsf{P}(H_{-m}) \\ &\le C + \pi_1(1-q) \sum_{m\ge 2} q^{2(m-1)} + (1-\pi_1) \sum_{m\ge 2} q^{3(m-1)} \\ &\le C + \pi_1(1-q)\frac{q^2}{1-q^2} + (1-\pi_1)\frac{q^3}{1-q^3}.\end{aligned}$$

This gives another upper bound on π_1 in terms of C, which, when combined with the trivial inequality (10.10), gives

$$C \geq (1-q^2)\frac{1+q-q^2-q^4+q^5}{1+2q-q^2-q^4+q^5-q^3+q^6} =: L_\infty(q). \tag{10.12}$$

Defining $L(q) = \max(L_1(q), L_\infty(q))$ gives the lower bound announced in Theorem 10.1.

We now turn to upper bounds. Note that, for all k,

$$\{X_0(0) = k, \xi_1 = k+1\} \subseteq \{X_1(0) = k+1\},$$

directly from the definition of Φ. This gives

$$\pi_k p q^k \leq \pi_{k+1}, \quad k \geq 1.$$

From this we have the bounds

$$\pi_k \geq \pi_1 p^{k-1} q^{k(k-1)/2}, \quad k \geq 2.$$

Using expression (10.7) for C, i.e., $C = 1 - \pi_1 q - \sum_{k\geq 2} \pi_k q^k$, we get

$$C \leq 1 - \pi_1 q - \pi_1 p f(q), \tag{10.13}$$

where $f(q) := \sum_{k\geq 1} p^{k-1} q^{(k+1)(k+2)/2}$. Now, consider the inclusion (10.4). We have

$$\mathsf{P}(H_{-m}) \geq \pi_1(1-q)q^2 q^3 \cdots q^m = \pi_1 p q^{(m-1)(m+2)/2}.$$

So, using (10.9), we obtain

$$\pi_1 \geq C + \pi_1 p g(q), \tag{10.14}$$

where $g(q) = \sum_{k\geq 1} q^{k(k+3)/2}$. Combining (10.13) with (10.14) we obtain

$$C \leq \frac{1-pg(q)}{1+q+pf(q)-pg(q)} =: U_1(q). \tag{10.15}$$

This proves the first part of the second inequality of Theorem 10.1.

Consider now the second auxiliary Lemma 10.2. Replace q with q^2 in (10.6) and substitute in itself to obtain

$$C(q) \leq \frac{1-q}{1+q} + \frac{q(1-q^2)}{(1+q)(1+q^2)} + \frac{q^3}{(1+q)(1+q^2)} C(q^4).$$

By induction, we get

$$C(q) \leq \sum_{n=0}^{\infty} \frac{q^{2^n-1}(1-q^{2^n})}{(1+q)(1+q^2)\cdots(1+q^{2^n})}.$$

The last denominator can be expressed as

$$\begin{aligned}(1+q)(1+q^2)\cdots(1+q^{2^n}) &= \frac{1-q^2}{1-q}\frac{1-q^4}{1-q^2}\cdots\frac{1-q^{2^{n+1}}}{1-q^{2^n}}\\ &= \frac{1-q^{2^{n+1}}}{1-q} = \frac{(1-q^{2^n})(1+q^{2^n})}{1-q},\end{aligned}$$

and so

$$C(q) \le (1-q)\sum_{n=0}^{\infty}\frac{q^{2^n-1}}{1+q^{2^n}} =: V(q). \tag{10.16}$$

So, our upper bound for C is $U(q) = \min\{U_1(q), V(q)\}$. □

Corollary 10.1.

$$C = \begin{cases} 1-q+q^2-3q^3+7q^4+O(q^5), & \text{as } q \to 0\\ O(p\log p), & \text{as } q \to 1.\end{cases}$$

Proof. Consider the functions $L_1(q)$, $L_\infty(q)$, $U_1(q)$, and $V(q)$, defined via equations (10.11), (10.12), (10.15), and (10.16), respectively, in the course of proof of Theorem 10.1. We also defined the functions $L(q) = \max(L_1(q), L_\infty(q))$, $U(q) = \min(U_1(q), V(q))$. It turns out that there is $q^* \in (0,1)$ such that $L_\infty(q) > L_1(q)$ for $0 < q < q^*$, and $L_\infty(q) < L_1(q)$ for $q^* < q < 1$. Using the symbolic calculator of the program MAPLE we compute the Taylor expansion of $L_\infty(q)$, given by (10.12), at $q = 0$:

$$L_\infty(q) = 1-q+q^2-3q^3+7q^4+O(q^5), \quad \text{as } q \to 0.$$

Similarly, it turns out that there is $q^{**} \in (0,1)$ such that $U_1(q) < V(q)$ for $0 < q < q^{**}$, while $U_1(q) > V(q)$ for $q^{**} < q < 1$. Using MAPLE, we computed the Taylor expansion of $U_1(q)$, given by (10.15), at $q = 0$:

$$U_1(q) = 1-q+q^2-3q^3+7q^4+O(q^5), \quad \text{as } q \to 0,$$

which is identical to that of $L_\infty(q)$. (It can be seen that terms of order ≥ 4 differ.) The two Taylor expansions match, and so this is the Taylor expansion for $C(q)$ at $q = 0$. To get asymptotics as $q \to 1$ (i.e., $p \to 0$), we use the bound $V(q)$. From its definition, we see that there is a constant c_2 such that

$$V(q) \sim c_2 p(q+q^2+q^4+q^8+\cdots).$$

The infinite sum in parentheses is asymptotically equivalent to $I(-\log q)$ as $p \to 0$, where

$$I(u) := \int_0^\infty \exp\{-u2^x\}\,dx.$$

It is easy to see that

$$I(u) := \frac{1}{\log 2} \int_u^\infty \frac{e^{-y}}{y}\, dy \sim \frac{\log u}{\log 2}, \qquad \text{as } u \to 0.$$

Setting $u = -\log q$ we get

$$I(-\log q) \sim \log_2 p, \qquad \text{as } p \to 0.$$

Hence

$$V(q) \sim c_2 p \log_2 p, \qquad \text{as } p \to 0.$$

□

Discussion

It is interesting to observe that $L(q)$ and $U(q)$ agree sharply for q up to about 0.2. Both functions $L(q)$ and $U(q)$ strictly decrease as q increases. See Figure 2 for a plot of these two functions. Observe that $U(q) \le 1/(1+q)$, which is already a good upper bound for small values of q. This bound is elementary: it follows from the inequalities $C \le 1 - \pi_1 q$, and $\pi_1 \ge C$.

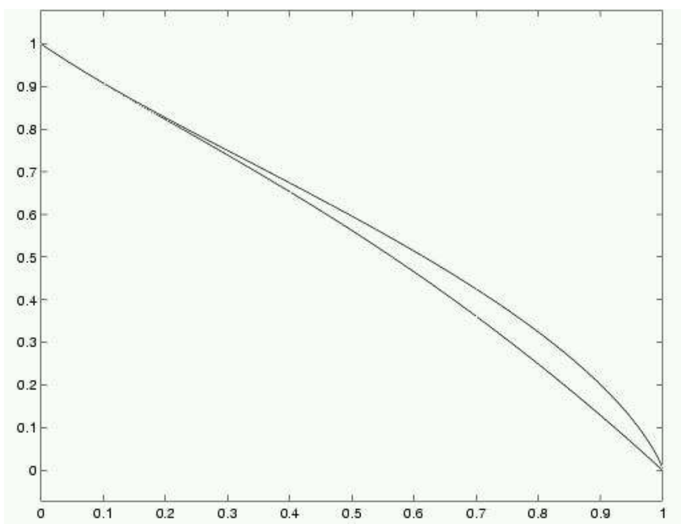

Figure 2. Upper and lower bounds on C as a function of $q = 1 - p$.

We also remark, without proof, that if we define the truncated series $f_n(q) = \sum_{k=1}^n p^{k-1} q^{(k+1)(k+2)/2}$, $g_n(q) = \sum_{k=1}^n q^{k(k+3)/2}$, then the functions $U_n(q) =$

$\big(1 - pg_n(q)\big)/\big(1 + q + pf_n(q) - pg_n(q)\big)$, $n = 1, 2, \ldots$ are all upper bounds to C, and $1/(1+q) > U_n(q) > U_{n+1}(q) > U(q)$, for all $n \geq 1$. As for the constant c_2 that appears in the upper bound $V(q)$, it is easily seen that $1/2 < c_2 < 1$. This upper bound was proved to be of order $O(p \log p)$ as $p \to 0$. However, the lower bound does not match this asymptotics. We do, however, believe that the upper bound is a better approximation to C for very small values of p.

11. Queueing applications, extensions, and future work

Consider customers entering a queueing facility at the epochs T_n of a renewal process. The customer that arrives at T_n has service time σ_n and also has a list of customers of indices $i < n$ that influence him, in the sense that the arriving customer has to wait until the last of the customers in his list leaves the system; at that time, he may start service which has duration σ_n. In other words, this is a model of a queue with precedence constraints. Let $\mathcal{L}_n$ be the list of customers influencing customer n; $\mathcal{L}_n$ is a subset of $\{n-1, n-2, \ldots\}$. Let b_n denote the time at which customer n begins his service, and e_n the time at which he ends. To obtain a recursion, notice first that

$$e_n = b_n + \sigma_n,$$

and

$$b_n = \max(T_n, \max_{j \in \mathcal{L}_n} e_j).$$

As an example, if the lists $\mathcal{L}_n$ are all empty then there are no constraints, and the system is simply a $G/G/\infty$ queue. On the other extreme, if $\mathcal{L}_n = \{n-1\}$, then the system is a First-Come First-Serve $G/G/1$ queue. The system satisfies the monotone separable framework of the saturation rule; see [4]. Thus, under stationary-ergodic assumptions, the queueing system is stable if

$$C < \mathsf{E}\tau,$$

where τ is a random variable that has the distribution of $T_{n+1} - T_n$ (given that $T_0 = 0$), and C is the constant defined as follows. Consider the above recursions but with $T_n = 0$ for all n and let $\tilde{e}_n$ be the corresponding end-of-service variables. Then, by the subadditive ergodic theorem,

$$C = \lim_{n \to \infty} n^{-1} \max_{i \leq n} \tilde{e}_i,$$

almost surely.

In Section 10, we have studied the case where all service times are equal to 1, and the $\mathcal{L}_n$ is chosen by tossing a coin with success probability p. Then the constant C just defined is the constant derived in Section 10.

More generally, assuming σ_n random, is a problem equivalent to studying longest paths in a random graph with weights, where "length" of a path is the sum of the weights of its edges. The problem of finding regeneration events that enable us to construct a stationary regime (which leads to the bounds for C) is, at the present moment, the subject of a forthcoming paper.

A. Appendix. Auxiliary results

Lemma A.1. *If Y_1, Y_2 and Z are three random variables such that Z is independent of (Y_1, Y_2), $\mathsf{P}(Z \neq 0) > 0$, and $Y_1 Z = Y_2 Z$, a.s., then $Y_1 = Y_2$, a.s.*

Proof. Since $\mathsf{P}(Y_1 Z = Y_2 Z) = 1$, we have

$$\begin{aligned}\mathsf{P}(Z \neq 0) = \mathsf{P}(Y_1 Z = Y_2 Z, Z \neq 0) &= \mathsf{P}(Y_1 = Y_2, Z \neq 0) \\ &= \mathsf{P}(Y_1 = Y_2)\mathsf{P}(Z \neq 0),\end{aligned}$$

where the last equality follows from independence. Since $\mathsf{P}(Z \neq 0) > 0$, we obtain the result $\mathsf{P}(Y_1 = Y_2) = 1$. $\square$

Proposition A.1. *Let $X_1, X_2, \dots$ be a stationary-ergodic sequence of random variables with $\mathsf{E}X_1^+ < \infty$. Then*

$$\frac{1}{n} \max_{1 \le i \le n} X_i \to 0, \qquad \textit{a.s. and in } L^1.$$

Proof. Without loss of generality, assume $X_n \geq 0$, a.s. Put $Y_n = \max_{1 \le i \le n} X_i$. Clearly,

$$Y_{n+k} \le \max_{1 \le i \le n} X_i + \max_{n+1 \le i \le n+k} X_i = Y_n + Y_k \circ \theta^n.$$

Kingman's subadditive ergodic theorem shows that $Y_n/n \to c$, a.s., where $c \geq 0$. We will show that $c = 0$. If $c > 0$ then, for any $0 < \varepsilon < c/2$, there is k_0 such that $\mathsf{P}(Y_k/k > c + \varepsilon) < \varepsilon$ for all $k \geq k_0$. Fix $k \geq k_0$ and let $n = 2k$. We then have

$$\begin{aligned}\mathsf{P}(Y_n/n > 3c/4) &\le \mathsf{P}(Y_n/n > (c+\varepsilon)/2) = \mathsf{P}(\max(Y_k/n, Y_k \circ \theta^k/n) > (c+\varepsilon)/2) \\ &\le 2\mathsf{P}(Y_k/n > (c+\varepsilon)/2) = 2\mathsf{P}(Y_k/k > c+\varepsilon),\end{aligned}$$

which contradicts the a.s. convergence of Y_n/n to c. Hence $c = 0$. To show that $\mathsf{E}Y_n/n \to 0$ simply observe that the sequence Y_n/n is bounded by $S_n/n = (X_1 + \cdots + X_n)/n$, and since $S_n/n \to \mathsf{E}X_1$, a.s. and in L^1, it follows that $\{S_n/n\}$ is a uniformly integrable sequence, and thus so is $\{Y_n/n\}$. $\square$

Proposition A.2. *Let $X_1, X_2, \dots$ be a stationary-ergodic sequence of random variables with $\mathsf{E}X_1 = 0$. Consider the stationary walk $S_n = X_1 + \cdots + X_n$, $n \geq 1$, with $S_0 = 0$. Put $M_n = \max_{0 \le i \le n} S_i$. Then $M_n/n \to 0$, a.s. and in L^1.*

Proof. Fix $\varepsilon > 0$. Let $M^* = \sup_{i \geq 0}(S_i - i\varepsilon)$. Then $M^* < \infty$, a.s. We have

$$M_n = \max_{0 \leq i \leq n} (S_i - i\varepsilon + i\varepsilon) \leq \max_{0 \leq i \leq n} (S_i - i\varepsilon) + n\varepsilon \leq M^* + n\varepsilon.$$

So, $M_n/n \leq M^*/n + \varepsilon$, a.s. This implies that $\limsup_{n\to\infty} M_n/n \leq \varepsilon$, a.s., and so $M_n/n \to 0$, a.s. Convergence in L^1 can be proved as follows. Let $b := \mathsf{E}X_1^+$. Define $S_n^{(+)} := X_1^+ + \cdots + X_n^+$. By the ergodic theorem, $S_n^{(+)}/n \to b$, a.s. and in L^1, and so $\{S_n^{(+)}/n\}$ is a uniformly integrable sequence. But $0 \leq M_n/n \leq S_n^{(+)}/n$. So $\{M_n/n\}$ is also uniformly integrable. □

References

[1] V. Anantharam and T. Konstantopoulos (1997) Stationary solutions of stochastic recursions describing discrete event systems. *Stoch. Process. Appl.* **68**, 181–194.

[2] V. Anantharam and T. Konstantopoulos (1999) A correction and some additional remarks on: "Stationary solutions of stochastic recursions describing discrete event systems". *Stoch. Process. Appl.* **80**, 271–278.

[3] F. Baccelli and P. Brémaud (1994) *Elements of Queueing Theory.* Springer-Verlag, Berlin.

[4] F. Baccelli and S. Foss (1995) On the saturation rule for the stability of queues. *J. Appl. Probab.* **32** (2), 494–507.

[5] P. Billingsley (1968) *Convergence of Probability Measures.* Wiley, New York.

[6] A.A. Borovkov (1978) Ergodicity and stability theorems for a class of stochastic equations and their applications. *Theory Probab. and Appl.* **23**, 227–258.

[7] A.A. Borovkov (1980) *Asymptotic Methods in Queueing Theory.* Nauka, Moscow (Russian edition). Revised English edition: Wiley, New York, 1984.

[8] A.A. Borovkov (1998) *Ergodicity and Stability of Stochastic Processes.* Wiley, Chichester.

[9] A.A. Borovkov and S.G. Foss (1992) Stochastically recursive sequences and their generalizations. *Sib. Adv. Math.* **2** (1), 16–81.

[10] A.A. Borovkov and S.G. Foss (1994) Two ergodicity criteria for stochastically recursive sequences. *Acta Appl. Math.* **34**, 125–134.

[11] A. Brandt, P. Franken and B. Lisek (1990) *Stationary Stochastic Models.* Akademie-Verlag, Berlin.

[12] P. Brémaud and L. Massoulié (1994) Imbedded construction of stationary sequences and point processes with a random memory. *Queueing Systems, Theory Appl.* **17**, 213–234.

[13] J.E. Cohen and C.M. Newman (1991) Community area and food chain length: theoretical predictions. *Amer. Naturalist* **138**, 1542–1554.

[14] F. Comets, R. Fernández and P.A. Ferrari (2001). Processes with long memory: regenerative construction and perfect simulation. Preprint.

[15] P. Diaconis and D. Freedman (1999) Iterated random functions. *SIAM Review* **41**, 45–76.

[16] S. Foss (1983) On ergodicity conditions in multi-server queues. *Sib. Math. J.* **34** (6), 168–175.

[17] S. Foss and R.L. Tweedie (1998) Perfect simulation and backward coupling. *Stoch. Models* **14** (1–2), 187–204.

[18] M. Isopi and C.M. Newman (1994) Speed of parallel processing for random task graphs. *Comm. Pure and Appl. Math.* **47**, 261–276.

[19] T. Lindvall *Lectures on the Coupling Method.* Wiley, New York.

[20] C.M. Newman (1992) Chain lengths in certain random directed graphs. *Random Struct. and Algorithms* **3** (3), 243–253.

[21] C.M. Newman and J.E. Cohen (1986) A stochastic theory of community food webs: IV. Theory of food chains in large webs. *Proc. Royal. Soc.* (*London*), *Ser. B* **228**, 355–377.

[22] J.G. Propp and D.B. Wilson (1996) Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Struct. and Algorithms* **9**, 223–252.

[23] H. Thorisson (2000) *Coupling, Stationarity, and Regeneration.* Springer-Verlag, New York.