PROCESSES WITH LONG MEMORY: REGENERATIVE CONSTRUCTION AND PERFECT SIMULATION

BY FRANCIS COMETS,[1] ROBERTO FERNÁNDEZ[2] AND PABLO A. FERRARI[2]

Université de Paris 7, Université de Rouen and Universidade de São Paulo

We present a perfect simulation algorithm for stationary processes indexed by $\mathbb{Z}$, with summable memory decay. Depending on the decay, we construct the process on finite or semi-infinite intervals, explicitly from an i.i.d. uniform sequence. Even though the process has infinite memory, its value at time 0 depends only on a finite, but random, number of these uniform variables. The algorithm is based on a recent regenerative construction of these measures by Ferrari, Maass, Martínez and Ney. As applications, we discuss the perfect simulation of binary autoregressions and Markov chains on the unit interval.

1. Introduction. In this paper we consider processes with transition probabilities that depend on the whole past history, that is, processes with long memory. When this dependence decays fast enough with time, we exhibit a regenerative construction which, besides yielding an explicit proof of existence and uniqueness of the process, can be transcribed into a perfect simulation scheme.

Processes with long memory have a long history. They were first studied by Onicescu and Mihoc (1935a, b) under the label *chains with complete connections* (*chaînes à liaisons complètes*). Harris (1955) proposed the somehow less used name of *chains of infinite order*. Doeblin and Fortet (1937) proved the first results on speed of convergence towards the invariant measure. Harris (1955) extended results on existence and uniqueness. The chains appeared also as part of the formalism introduced by Keane (1971, 1976) to study of subshifts of finite type (or covering transformations). In this theory the transition probabilities are called *g-functions* and the invariant measures *g-measures*.

The theory of long-memory processes has found applications in the study of urn schemes [Onicescu and Mihoc (1935b)], continued-fraction expansions [Doeblin (1940); Iosifescu (1978) and references therein], learning processes [see, e.g., Iosifescu and Theodorescu (1969), Norman (1972, 1974)], models of gene population [Norman (1975)], image coding [Barnsley, Demko, Elton and Gerinomo (1988)], automata theory [Herkenrath and Theodorescu (1978)],

partially observed—or "grouped"—random chains [Harris (1955), Blackwell (1957), Kaijser (1975), Pruscha and Theodorescu (1977), Elton and Piccioni (1992)] and products of random matrices [Kaijser (1981)]. For further references we refer the reader to Kaijser (1981, 1994), from which most of the material of the previous review paragraphs is taken, and to Iosifescu and Grigorescu (1990).

It is clear that these applications should benefit from the construction and perfect simulation scheme presented here. As an illustration, we discuss in Sections 3 and 9 applications to binary autoregressions and to the Markov processes on the interval $[0, 1]$ defined by Harris (1955) by mapping chains with complete connection into D-ary expansions.

In this paper we rely on a regenerative construction of the chain, which generalizes, in some direction, those existing in the literature. This type of construction was first introduced by Doeblin (1938) for Markov chains with countable alphabet. Schemes for more general state spaces came much later [Athreya and Ney (1978), Nummelin (1978)]. The first regenerative structures for chains with complete connections were proposed by Lalley (1986, 2000) and Berbee (1987) for chains with summable continuity rates. An explicit regenerative construction was put forward by Ferrari, Maass, Martínez and Ney (2000) in the spirit of Berbee's approach.

In the present paper we take up the scheme of Ferrari, Maass, Martínez and Ney (2000), extend it to part of the Harris uniqueness regime and transcribe it as a perfect simulation algorithm. Basically, the construction used here can be interpreted as a simultaneous coupling of *all* histories, built in such a way that at each instant i there is a (random) number $k_i \geq 0$ such that the distribution of the move $i + 1$ is the same for all histories agreeing the k_i preceding instants. This independence from the k_i-remote past yields the times τ such that preceding histories are irrelevant for future moves. These are the regeneration times, defined by the conditions $k_i \leq i - \tau$ for all $i \geq \tau$. Both Berbee's and Lalley's constructions rely on the regeneration probability being positive, a fact that seems to hold only for summable continuity rates. In contrast, our construction extends to cases where (global) regeneration may have probability zero. Non-summable—but still not too slowly decreasing—rates inside the Harris uniqueness regime yield *local* regeneration times, that is, regenerations for finite time intervals (windows). In ergodic-theory terms, our construction is in fact a finitary coding of a process of i.i.d. uniform variables in the interval $[0, 1]$.

Perfect simulation became popular after Propp and Wilson (1996) introduced the *coupling from the past* algorithm to simulate invariant measures of Markov chains. Wilson's page (http://dimacs.rutgers.edu/~dbwilson/exact) provides updated and extensive references on perfect simulation. Foss and Tweedie (1998) and Corcoran and Tweedie (2001) proposed a general framework based on regeneration schemes for Markov chains and the so-called "stochastic recursive sequences." These are processes defined by $X_{n+1} = f(X_n, \xi_n)$, where ξ is a stationary process. In the Markovian case, ξ_i are i.i.d.; in the non-Markovian case, it

remains the matter of how to construct or simulate the sequence ξ_i. Our algorithm applies to a wide variety of non-Markovian processes and, through the formalism of random systems with complete connections, it can be used to simulate Markov processes with state space of large cardinality (e.g., the unit interval). Section 9 present an example along these lines.

Our main result is an explicit construction, as a deterministic function of a sequence of i.i.d. random variables uniformly distributed in [0, 1], of realizations of the stationary chain with infinite memory. As corollaries we get (1) an alternative proof of the existence and uniqueness of the stationary process, and (2) a perfect simulation algorithm and a regeneration scheme for this process. These results are summarized in Theorem 4.1 and Corollaries 4.1, 4.2 and 4.3. In Section 2 we introduce the basic definitions and in Section 3 we illustrate the simulation on a concrete example, which is continued in Section 9. The results of Section 4 are proved in Sections 5, 6 and 8, though the perfect simulation algorithm is described in Section 7 in its general version.

2. Definitions. We denote by G our alphabet, $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ and $-\mathbb{N}^* = \{-i : i \in \mathbb{N}^*\}$. In what follows G can be finite or countable, though in the latter case conditions (4.3) and (4.8) below impose severe limitations. The set $G^{-\mathbb{N}^*}$ is the space of *histories*; we write $\omega_a^b := (w_b, w_{b-1}, \ldots, w_a)$ for $-\infty \le a \le b \le +\infty$. For brevity we write $\underline{w} = \omega_{-\infty}^\infty$. Let $P : G \times G^{-\mathbb{N}^*} \to [0, 1]$ be a probability transition kernel; that is, $P(g|w_{-\infty}^{-1}) \ge 0$ for all g and

$$(2.1) \qquad \sum_{g \in G} P(g|w_{-\infty}^{-1}) = 1$$

for each $w_{-\infty}^{-1} \in G^{-\mathbb{N}^*}$. The kernel P defines, by telescopic products, what in statistical mechanics is called a *specification*. A specification is a consistent system of conditional probabilities, where consistency is required for *all* histories $w_{-\infty}^{-1}$. In standard probabilistic treatments, such requirements are made only almost surely with respect to some preestablished appropriate measure. However, in the present setting the determination of the appropriate measure is part of the problem, and stronger requirements are a priori necessary.

Denoting by $\eta(i)$, $i \in \mathbb{Z}$, the coordinate mappings on $G^\mathbb{Z}$, we say that a (nonnecessarily stationary) probability measure ν on $G^\mathbb{Z}$—or a *process* with distribution ν—is *compatible* with the specification P if the latter is a version of the one-sided conditional probabilities of the former,

$$(2.2) \qquad \nu\big(\eta : \eta(i) = g \mid \eta(i + j) = w_j, \ j \in -\mathbb{N}^*\big) = P(g|w_{-\infty}^{-1})$$

for all $i \in \mathbb{Z}$, $g \in G$ and ν-a.e. $w_{-\infty}^{-1} \in G^{-N^*}$. Then the identities

$$(2.3) \qquad \begin{aligned} &\nu\big(\eta : \eta(i + k) = g_k, k \in [0, n] \mid \eta(i + j) = w_j, \ j \in -\mathbb{N}^*\big) \\ &\qquad = \prod_{k=0}^n P(g_k|g_{k-1}, \ldots, g_0, w_{-\infty}^{-1}), \end{aligned}$$

hold for ν-a.e. $w_{-\infty}^{-1}$, where the concatenation is defined by

$$(w_k, w_{k-1}, \ldots, w_\ell, z_{-\infty}^{-1}) := (w_k, w_{k-1}, \ldots, w_\ell, z_{-1}, z_{-2}, \ldots).$$

The regenerative construction of this paper is based on the functions

(2.4)
$$a_0(g) := \inf\{P(g|z_{-\infty}^{-1}) : z_{-\infty}^{-1} \in G^{-\mathbb{N}^*}\},$$

$$a_k(g|w_{-k}^{-1}) := \inf\{P(g|w_{-1}, \ldots, w_{-k}, z_{-\infty}^{-1}) : z_{-\infty}^{-1} \in G^{-\mathbb{N}^*}\}, \qquad k \geq 1,$$

defined for $k \in \mathbb{N}$, $g, w_{-1}, \ldots, w_{-k} \in G$. [These functions are denoted $g(i_0|i_{-1}, \ldots, i_{-k})$ by Berbee (1987).] The numbers

(2.5)
$$a_k := \inf_{w_{-k}^{-1}} \sum_{g \in G} a_k(g|w_{-k}^{-1}),$$

$k \in \mathbb{N}$, determine a probabilistic threshold for memories limited to k preceding instants. The sequences $a_k(g|w_{-k}^{-1})$ and a_k are nondecreasing in k and contained in $[0, 1]$.

For our construction we shall use a sequence $\underline{U} = (U_i : i \in \mathbb{Z})$ of independent random variables with uniform distribution in $[0, 1[$, constructed on the corresponding canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote $\mathbb{E}$ the expectation with respect to $\mathbb{P}$.

3. An example: binary autoregressive processes.

To motivate the method, we present an example that shows how to construct and perfectly simulate a process with infinite memory. Let us consider binary autoregressive processes. Such a process is the binary version of autoregressive (long memory) processes used in statistics and econometrics. It describes binary responses when covariates are historical values of the process [McCullagh and Nelder (1989), Section 4.3].

Let the state space be $G = \{-1, +1\}$, θ_0 a real number and $(\theta_k; k \geq 1)$ a summable real sequence. Let $q : \mathbb{R} \mapsto]0, 1[$ be strictly increasing and continuously differentiable. Assume that $P(\cdot|w_{-\infty}^{-1})$ is the Bernoulli law on $\{-1, +1\}$ with parameter

(3.1)
$$q\left(\theta_0 + \sum_{k \geq 1} \theta_k w_{-k}\right),$$

that is, $P(+1|w_{-\infty}^{-1}) = q(\theta_0 + \sum_{k \geq 1} \theta_k w_{-k}) = 1 - P(-1|w_{-\infty}^{-1})$.

By compactness there exists at least one process compatible with (3.1); the conditions for uniqueness are well known. The question is how to simulate (construct) such a stationary process. To do that we construct a family of partitions of the interval $[0, 1[$ indexed by k and w_{-k}^{-1} using $a_k(1|w_{-k}^{-1})$ and a_k defined in the previous section. Letting

$$r_k = \sum_{m > k} |\theta_m|$$

we have

$$a_k(1|w_{-k}^{-1}) = q\left(\theta_0 + \sum_{1 \le m \le k} \theta_m w_{-m} - r_k\right),$$

(3.2)

$$a_k(-1|w_{-k}^{-1}) = 1 - q\left(\theta_0 + \sum_{1 \le m \le k} \theta_m w_{-m} + r_k\right)$$

and

$$a_k = 1 - \sup_{w_{-k}^{-1}}\left\{q\left(\theta_0 + \sum_{1 \le m \le k} \theta_m w_{-m} + r_k\right)\right.$$

(3.3)

$$\left. - q\left(\theta_0 + \sum_{1 \le m \le k} \theta_m w_{-m} - r_k\right)\right\}.$$

For each k and w_{-k}^{-1}, let $\mathbf{B}_k(g|w_{-k}^{-1})$ be intervals of length $a_k(g|w_{-k}^{-1}) - a_{k-1}(g|w_{-k+1}^{-1})$ placed consecutively in lexicographic order in k and g, starting at the origin. Use the conventions $a_0(g|w_0^{-1}) = a_0(g)$ and $a_{-1} = 0$. Since $a_k \nearrow 1$ as k increases to infinity, these intervals form a partition of the interval $[0, 1[$.

Let $(U_i : i \in \mathbb{Z})$ be the sequence of i.i.d. random variables uniformly distributed in $[0, 1[$ defined at the end of Section 2. For $n \in \mathbb{Z}$ define the random variable

(3.4) $$K_n := \sum_{k \ge 0} k \, \mathbf{1}\{U_n \in [a_{k-1}, a_k)\}.$$

In our construction the variable $K_n := K_n(\underline{U})$ indicates how many sites in the past are needed to compute the state at time n. To each n associate an arrow going from n to $n - K_n$. The state at site n will be independent of the states at $t \le s$ if no arrow starting at $\{s, \ldots, n\}$ finishes to the left of s. Let $\tau[n] := \tau[n](\underline{U})$ be the maximum of such s,

(3.5) $$\tau[n] := \max\{s \le n : U_j < a_{j-s}, \ j \in [s, n]\}$$

$$= \max\{s \le n : K_j \ge s, \ j \in [s, n]\}.$$

Notice that $\tau[n]$ is a stopping time for the sequence $(U_{n-k} : k \ge 0)$. We show in Theorem 4.1 that the condition

(3.6) $$\sum_k r_k < \infty$$

is sufficient to guarantee $\mathbb{P}(\tau[n] > -\infty) = 1$ which, in turn, is an equivalent condition to the feasibility of the following construction.

Simulation (construction) of the stationary measure.

1. Generate successively i.i.d. uniform random variables $U_n, U_{n-1}, \ldots$. Stop when $U_{\tau[n]}$ is generated. Using (3.4), the values $K_n, K_{n-1}, \ldots, K_{\tau[n]}$ are simultaneously obtained.
2. Use the $U_{\tau[n]}, \ldots, U_n$ and $K_{\tau[n]}, \ldots, K_n$ generated in the previous step to define

$$(3.7) \qquad X_j = g \qquad \text{if } U_j \in \bigcup_{\ell=0}^{K_j} \mathbf{B}_\ell(g | X_{j-\ell}^{j-1}),$$

recursively from $j = \tau[n]$ to $j = n$.
3. Return X_n. The algorithm has also constructed $X_{\tau[n]}, \ldots, X_{n-1}$.

The expression (3.7) is well defined because by the definition of $\tau[n]$, $j - K_j \geq \tau[n]$, $U_j \in [0, a_{K_j}[\subset \bigcup_{g \in G} \bigcup_{\ell=0}^{K_j} \mathbf{B}_\ell(g | X_{j-\ell}^{j-1})$ and the set in (3.7) depends at most on $X_{j-K_j}, \ldots, X_{j-1}$. This is discussed in detail in Section 6; see (6.3). We show in Theorem 4.1 that the above algorithm constructs a realization X_n of a random variable which has the one-coordinate marginal of the unique measure compatible with the specification. To construct a realization of the measure in a finite window, just repeat the algorithm for other n reusing always previously generated U_j. The algorithm induces a function $\Phi : [0, 1[^{\mathbb{Z}} \to \{-1, 1\}^{\mathbb{Z}}$. Given the event $\{\tau[n] = k\}$, X_n—the nth coordinate of $\Phi(U_{-\infty}^{\infty})$—depends only on $U_k, \ldots, U_n$. (To be rigorous, one should give the definition of Φ when $\tau[n] = -\infty$ for some n; this is an arbitrary but irrelevant choice as this set has probability zero under our hypotheses.)

This construction exploits a well-known fact. The existence of a renovating event gives rise to a perfect simulation algorithm and a regenerative structure. In our case there is a regeneration at time s if $j - K_j \geq s$ for all $j \geq s$ (no arrow passes over s). However, for the construction of the measure in site n we use a weaker condition: it suffices that no arrow passes from j to the left of s for $j \in [s, n]$.

While in some cases one may have explicit expressions for a_k, in general, this will not be the case [see (3.3)]. A useful aspect of our construction is that in these cases we can work with *lower bounds* a_k^*. We shall discuss this issue in Section 9.

4. Results. The existence of our regeneration scheme depends on the nonincreasing sequence

$$(4.1) \qquad \beta_m := \prod_{k=0}^{m} a_k,$$

with a_k defined in (2.5), and a necessary (but not sufficient) condition is $\lim_{k \to \infty} a_k = 1$. The regeneration time for a window $[s, t]$, for $-\infty < s < \infty$ and $s \leq t \leq \infty$, is the random variable

$$(4.2) \qquad \tau[s, t] := \max\{m \leq s : U_k < a_{k-m}, k \in [m, t]\},$$

which may be $-\infty$. Notice that $\tau[s,t] = \min\{\tau[n]:n \in [s,t]\}$ and that it is a *stopping time* for $(U_{t-k}:k \geq 0)$, in the sense that $\{\tau[s,t] \geq j\} \in \mathcal{F}(U_i:i \in [j,t])$ for $j \leq s$, where $\mathcal{F}(U_i:i \in [j,t])$ is the sigma field generated by $(U_i:i \in [j,t])$.

Our main result is the following theorem.

THEOREM 4.1.   *If*

(4.3)
$$\sum_{m \geq 0} \beta_m = \infty,$$

then:

   (i) *For each finite* $[s,t] \subset \mathbb{Z}$, $\mathbb{P}(\tau[s,t] > -\infty) = 1$, *where* $\tau[s,t]$ *is defined in* (4.2).

   (ii) *There exists a measurable function* $\Phi:[0,1[^{\mathbb{Z}} \rightarrow G^{\mathbb{Z}}$, *described in Section* 6, *such that*

(4.4)
$$\mu := \mathbb{P}(\Phi(\underline{U}) \in \cdot),$$

the law of $\Phi(\underline{U})$, *is compatible with* P. *Moreover, the distribution* μ *is stationary.*

   (iii) *In addition, the function* Φ *has the property that for each finite interval* $[s,t] \subset \mathbb{Z}$, *its restriction* $\{\Phi(\underline{U})(i):i \in [s,t]\}$ *depends only on the values of* U_i *in the interval* $[\tau[s,t],t]$. *More precisely, for* $i \in [s,t]$,

(4.5)
$$\Phi(\underline{U})(i) = \Phi(\ldots, v_{\tau-2}, v_{\tau-1}, U_\tau, \ldots, U_t, v_{t+1}, \ldots)(i)$$

for any sequence $\underline{v} \in [0,1[^{\mathbb{Z}}$ *(we abbreviated* $\tau[s,t]$ *as* τ*).*

   (iv) *The law of* $\tau[s,t]$ *satisfies the following bound*:

(4.6)
$$\mathbb{P}(s - \tau[s,t] > m) \leq \sum_{i=0}^{t-s} \rho_{m+i},$$

where ρ_m *is the probability of return to the origin at epoch m of the Markov chain on* $\mathbb{N}$ *starting at time zero at the origin with transition probabilities*

(4.7)
$$p(x,x+1) = a_x,$$
$$p(x,0) = (1 - a_x)$$

and $p(x,y) = 0$ *otherwise. In particular, if* $(1 - a_k)$ *decreases exponentially fast with k, then so does* ρ_k. *If* $(1 - a_k) \sim k^{-\gamma}$ *with* $\gamma > 1$ *then* ρ_k *decreases with the same power.*

   (v) *If*

(4.8)
$$\beta := \lim_{m \to \infty} \beta_m > 0,$$

then items (i), (iii) *and* (iv) *above hold also for* $t = \infty$.

This theorem is proved in Sections 5 and 6. More detailed bounds on the parameters ρ_m are given in Proposition 5.1. Conditions (4.3) and (4.8) require $a_0 > 0$ as in Harris (1955). Both conditions are noticeable weaker than those imposed by Lalley (1986), Berbee (1987) and Bressaud, Fernández and Galves (1999), as well as those corresponding to the g-measure approach (Ledrappier, 1974) and to Gibbsian specifications (Kozlov, 1974).

REMARK 4.1.   The construction can be performed replacing the a_k's with lower bounds, that is, considering a sequence a_k^* in $]0, 1[$, such that

(4.9) $a_k^* \leq a_k$ for $k \geq 0$.

The theorem is valid replacing the unstarred a_k by starred ones, and using the corresponding starred versions of β_m, τ and ρ_m. While the actual a_k's give shorter regeneration times, they could be hard to estimate. Suitable choices of a_k^* could provide a reasonable compromise between shorter regeneration times and feasible calculational efforts (Section 9).

Our first corollary is the uniqueness of the measure compatible with P.

COROLLARY 4.1 (Loss of memory and uniqueness).   (i) *Every measure μ compatible with the specification $\mathbb{P}$ has the following loss-of-memory property*: *If f is a function depending on the interval $[s, t]$ and $i \leq s$,*

(4.10)
$$\left| \mu(f \mid \eta(j) = w_j, \ j < i) - \mu(f \mid \eta(j) = v_j, \ j < i) \right|$$
$$\leq 2\|f\|_\infty \sum_{j=0}^{t-s} \rho_{s+j-i}$$

for every $\underline{w}, \underline{v} \in G^{\mathbb{Z}}$. Here we use the notation $\mu f := \int \mu(d\eta) f(\eta)$.
   (ii) *If $\sum_{m \geq 0} \beta_m = \infty$ the measure μ defined in (4.4) is the unique measure compatible with P.*

The uniqueness result is not new. Under the more restrictive condition (4.8) it was already obtained by Doeblin and Fortet (1937). Harris (1955) [see also Section 5.5 of Iosifescu and Grigorescu (1990) and references therein] extended this uniqueness to a region that coincides with (4.3) for two-symbol alphabets but it is larger for larger alphabets. Other uniqueness results, in smaller regions, were obtained by Ledrappier (1974) and Berbee (1987) in different ways. Results on loss of memory were also obtained by Doeblin and Fortet (1937) [see also Iosifescu (1992)], under the summability condition (4.8). Bressaud, Fernández and Galves (1999) extended those to a region defined by a condition slightly stronger than (4.3). The rates of loss of memory obtained in this last references strengthen those of Doeblin, Fortet and Iosifescu, but are weaker than ours.

A corollary of (ii) and (iii) of the theorem is a perfect simulation scheme.

COROLLARY 4.2 (Perfect simulation). *Let P be a specification with $\sum_{m\geq 0}\beta_m = \infty$ and μ the unique measure compatible with P. For each finite window $[s,t]$ there exist a family $(\widetilde{\Phi}_{m,t}; m \leq s)$ of functions $\widetilde{\Phi}_{m,t}:[0,1[^{[m,t]} \to G^{[m,t]}$ such that*

$$(4.11) \quad \mu(\eta:\eta(i)=g_i,\ i\in[s,t]) = \mathbb{P}(\underline{U}:\widetilde{\Phi}_{\tau,t}(U_\tau,\ldots,U_t)(i)=g_i,\ i\in[s,t])$$

where $\tau = \tau[s,t]$ is the stopping time defined in (4.2).

Expression (4.11) is our perfect simulation scheme for μ. Possible implementations are discussed in Section 7. Roughly speaking the algorithm goes as follows:

1. Produce a realization of $U_t,\ldots,U_s,U_{s-1},\ldots,U_{\tau[s,t]}$.
2. Compute $\widetilde{\Phi}_{\tau[s,t],t}(U_{\tau[s,t]},\ldots,U_t)$.
3. Return the configuration $(\widetilde{\Phi}_{\tau[s,t],t}(U_{\tau[s,t]},\ldots,U_t)(i):i\in[s,t])$. Its law is the projection of μ in the window $[s,t]$.

Our algorithm shares two features with the famous coupling from the past (CFTP) algorithm for Markov chains, introduced by Propp and Wilson (1996): (1) The r.v. U_i are generated sequentially backwards in time, until the stopping time $\tau[s,t]$ is discovered. (2) The algorithm is based on a coupled realization of the chain for all possible initial conditions ensuring the coalescence of the different trajectories before the observation window. Restricted to the Markovian case, our renewal times are, in principle, larger than the coalescence times of well-designed CFTP algorithms. Nevertheless, our approach, besides extending to non-Markovian processes, yields coupling times that depend only on the numbers a_k defined in (2.5) [see (4.12)], and hence that are only indirectly related to the cardinality of the alphabet.

The bounds in (iv) of Theorem 4.1 can be used to control the *user-impatience bias*. This picturesque name, proposed by Fill (1998), relates to the bias caused by practical constraints such as time limitation or computer storage capacity which force the user to stop a (long and unlucky) run before it reaches perfect equilibrium. Indeed, suppose that while sampling window $[s,t]$ we decide to abort runs with regeneration times $\tau[s,t] > M$ for some large $M > 0$, causing our algorithm to produce a biased distribution $\widehat{\mu}_{[s,t]}^M$. Applying Lemma 6.1 of Fill (1998) and (4.6), we obtain:

$$(4.12) \qquad \sup_{A\in G^{[s,t]}}|\mu(A)-\widehat{\mu}_{[s,t]}^M(A)| \leq \frac{\sum_{i=0}^{t-s}\rho_{M+i}}{1-\sum_{i=0}^{t-s}\rho_{M+i}}.$$

In regime (v), Theorem 4.1 yields the following *regeneration* scheme. Let $\mathbf{N} \in \{0,1\}^{\mathbb{Z}}$ be the random counting measure defined by

$$(4.13) \qquad\qquad \mathbf{N}(j) := \mathbf{1}\{\tau[j,\infty]=j\}.$$

Let $(T_\ell : \ell \in \mathbb{Z})$ be the ordered time events of $\mathbf{N}$ defined by $\mathbf{N}(i) = 1$ if and only if $i = T_\ell$ for some ℓ, $T_\ell < T_{\ell+1}$ and $T_0 \leq 0 < T_1$.

COROLLARY 4.3 (Regeneration scheme). *If $\beta > 0$, then the process $\mathbf{N}$ defined in* (4.13) *is a stationary renewal process with renewal distribution*

$$(4.14) \qquad \mathbb{P}(T_{\ell+1} - T_\ell \geq m) = \rho_m$$

for $m > 0$ and $\ell \neq 0$. Furthermore, the random vectors $\xi_\ell \in \bigcup_{n \geq 1} G^n$, $\ell \in \mathbb{Z}$, defined by

$$(4.15) \qquad \xi_\ell(\underline{U}) := \big(\Phi(\underline{U})(T_\ell), \ldots, \Phi(\underline{U})(T_{\ell+1} - 1)\big)$$

are mutually independent and $(\xi_\ell(\underline{U}) : \ell \neq 0)$ are identically distributed.

5. Distribution of $\tau[s, t]$ and $\tau[s, \infty]$. In this section we prove items (i) and (iv) [and the corresponding parts in (v)] of the theorem. We fix an increasing sequence of numbers (a_k) such that

$$(5.1) \qquad a_k \nearrow 1 \qquad \text{as } k \nearrow \infty$$

and define the following *house of cards* family of chains $((W_n^m : n \geq m) : m \in \mathbb{Z})$ by $W_m^m = 0$ and for $n > m$,

$$(5.2) \qquad W_n^m := (W_{n-1}^m + 1)\mathbf{1}\{U_n < a_{W_{n-1}^m}\},$$

where U_n are the uniform random variables introduced in the previous section.

Notation alert. Here W_n^m is not a vector of lenght $n - m$, but the position at time n of a chain starting at time m at the origin. Alternatively, $\mathbb{P}(W_n^m = y \mid W_{n-1}^m = x) = p(x, y)$, where the latter are the transition probabilities defined in item (iv) of Theorem 4.1. Notice that

$$(5.3) \qquad \rho_k = \mathbb{P}(W_{m+k}^m = 0)$$

for all $m \in \mathbb{Z}$ and $k \in \mathbb{N}^*$. The monotonicity of the a_k's implies that

$$(5.4) \qquad W_n^m \geq W_n^k \qquad \text{for all } m < k \leq n.$$

Hence, $W_n^m = 0$ implies that $W_n^k = 0$ for $m < k \leq n$, and the chains *coalesce* at time n:

$$(5.5) \qquad W_n^m = 0 \quad \implies \quad W_t^m = W_t^k, \qquad m \leq k \leq n \leq t$$

(furthermore, $W_t^m > W_t^k$ for t smaller than the smallest such n).

By the definition (4.2) of τ, we have for $j \leq s$,

$$
\begin{aligned}
\tau[s, t] < j \quad &\Longleftrightarrow \quad \forall m \in [j, s], \ \exists n \in [m, t] : W_n^{m-1} = 0, \\
&\Longleftrightarrow \quad \forall m \in [j, s], \ \exists n \in [s, t] : W_n^{m-1} = 0, \\
(5.6) \\
&\Longleftrightarrow \quad \max\{m < s : \forall n \in [s, t], \ W_n^m > 0\} < j - 1, \\
&\Longleftrightarrow \quad \exists n \in [s, t] : W_n^{j-1} = 0,
\end{aligned}
$$

where the second line follows from the coalescing property (5.5), and the last line is a consequence of the second one and the monotonicity (5.4). From the third line we get

$$(5.7) \qquad \tau[s, t] = 1 + \max\{m < s : W_n^m > 0, \ \forall n \in [s, t]\}.$$

LEMMA 5.1. (a) *Condition* $\sum_{n \geq 0} \prod_{k=1}^n a_k = \infty$ *is equivalent to any of the two following properties*:

(a.1) *The chain* $(W_n^m : n \geq m)$ *is nonpositive-recurrent for each* m.
(a.2) *For each* $-\infty < s \leq t < \infty$, $\tau[s, t] > -\infty$ *a.s.*

(b) *Condition* $\prod_{k=1}^\infty a_k > 0$ *is equivalent to any of the two following properties*:

(b.1) $(W_n^m : n \geq m)$ *is transient.*
(b.2) *For each* $-\infty < s$, $\tau[s, \infty] > -\infty$ *a.s.*

In both cases, *inequality* (4.6) *holds*.

PROOF. (a) It is well known that W_n is positive-recurrent if and only if $\sum_{n \geq 0} \prod_{k=1}^n a_k < \infty$ [e.g., Dacunha-Castelle, Duflo and Genon-Catalot (1983), page 62, Example E.4.1]. Also, by (5.6),

$$(5.8) \qquad \{\tau[s, t] < m\} = \bigcup_{i \in [s,t]} \{W_i^{m-1} = 0\}$$

for $m \leq s$. By translation invariance, the probability of the right-hand side of (5.8) satisfies

$$(5.9) \qquad \mathbb{P}\left(\bigcup_{i \in [s,t]} \{W_{-m+i}^0 = 0\}\right) \in \left[\mathbb{P}(W_{t-m}^0 = 0), \sum_{i=1}^{t-s} \mathbb{P}(W_{s-m+i}^0 = 0)\right]$$

where the right-hand side is a consequence of the monotonicity property (5.4). As $m \to -\infty$ this interval remains bounded away from 0 in the positive-recurrent case, but shrinks to 0 otherwise. Since $\mathbb{P}(\tau[s, t] = -\infty) = \lim_{m \to -\infty} \mathbb{P}(\tau[s, t] < m)$, this proves that (a.1) and (a.2) are equivalent.

(b) Clearly,

$$(5.10) \qquad \mathbb{P}(W_n^m \neq 0, \ \forall n \geq m) = \prod_{i=0}^\infty a_i.$$

This implies that the product of a_k is positive if and only if the house-of-cards process is transient. From (5.8) we have that

$$(5.11) \qquad \mathbb{P}(\tau[s, \infty] < m) = \mathbb{P}\left(\bigcup_{i \in [s-m+1, \infty]} \{W_i^0 = 0\}\right),$$

which goes to zero as $m \to -\infty$ in the transient case only. Therefore (b.1) and (b.2) are equivalent.

Inequality (4.6) follows from (5.8) or (5.11), due to (5.3). $\quad\square$

The following proposition is due to Bressaud, Fernández and Galves (1999).

PROPOSITION 5.1. *Let a_k be a $[0, 1]$-sequence increasing to one. Let ρ_k be the probability of return to the origin at epoch k of $(W_n^0 : n \geq 0)$.*

  (i) *If $\sum_{n\geq 0} \prod_{k=1}^n a_k = \infty$, then $\rho_n \to 0$.*
  (ii) *If $\prod_{k=1}^\infty a_k > 0$, then $\sum_{n\geq 0} \rho_n < \infty$.*
  (iii) *If $(1 - a_n)$ decreases exponentially then so does ρ_n.*
  (iv)

$$(5.12) \quad \prod_{k=1}^\infty a_k > 0 \quad and \quad \sup_i \limsup_{k\to\infty} \left(\frac{1 - a_i}{1 - a_{ki}}\right)^{1/k} \leq 1 \quad imply \quad \rho_n = \mathrm{O}(1 - a_n).$$

Item (iv) can be applied, for instance, when $a_n \sim 1 - (\log n)^b n^{-\gamma}$ for $\gamma > 1$. Items (i) and (ii) are direct transcriptions of (a.1) and (b.2) of the previous lemma.

REMARK 5.1. The results of this section are also monotonic on the choice of sequence (a_k): A sequence (a_k^*), satisfying (5.1), with $a_k^* \leq a_k$ for all k, yields chains W_n^{*m} with larger probability of visiting 0, and hence with larger regeneration intervals $s - \tau^*[s, t]$ and larger values of ρ_m^*. This is the content of Remark 4.1.

6. Construction of Φ. In this section we prove results needed to show (ii) and (iii) [and the corresponding part in (v)] of Theorem 4.1. The results hold for any sequence (a_k^*) satisfying (4.9), but for notational simplicity we omit the superscript "*". For $g \in G$ let $b_0(g) := a_0(g)$ and for $k \geq 1$,

$$b_k(g|w_{-k}^{-1}) := a_k(g|w_{-k}^{-1}) - a_{k-1}(g|w_{-k}^{-1}).$$

For each $w_{-\infty}^{-1} \in G^{-\mathbb{N}^*}$ let $\{\mathbf{B}_k(g|w_{-k}^{-1}) : g \in G, \ k \in \mathbb{N}\}$ be a partition of $[0, 1[$ with the following properties: (1) for $g \in G$, $k \geq 0$, $\mathbf{B}_k(g|w_{-k}^{-1})$ is an interval closed in the left extreme and open in the right one with Lebesgue measure $|\mathbf{B}_k(g|w_{-k}^{-1})| = b_k(g|w_{-k}^{-1})$; (2) these intervals are disposed in increasing lexicographic order with respect to g and k in such a way that the left extreme of one interval coincides with the right extreme of the precedent,

$$\mathbf{B}_0(g_1), \mathbf{B}_0(g_2), \ldots, \mathbf{B}_1(g_1|w_{-1}^{-1}), \mathbf{B}_1(g_2|w_{-1}^{-1}), \ldots, \mathbf{B}_2(g_1|w_{-2}^{-1}), \mathbf{B}_2(g_2|w_{-2}^{-1}), \ldots$$

is a partition of $[0,1[$ into consecutive intervals increasingly arranged. This definition is illustrated in Figure 1 in the case $G = \{1, 2\}$.
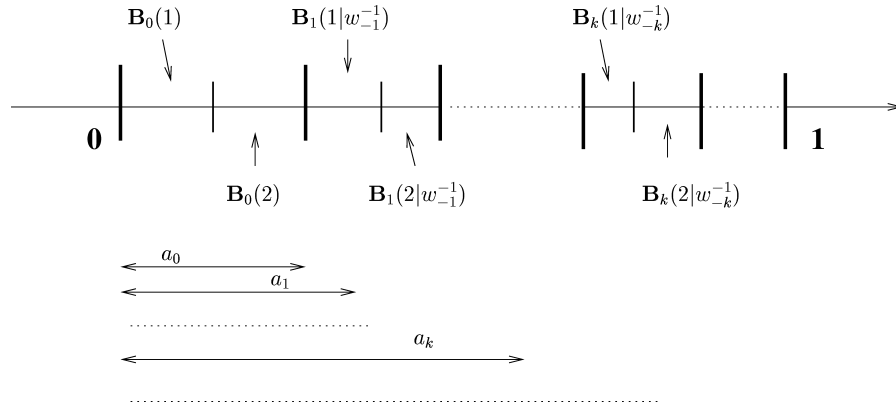
FIG. 1. *Partition* $\{\mathbf{B}_k(g|w_{-k}^{-1}) : g \in G, \ k \in \mathbb{N}\}$ *of* $[0, 1[$ *in the case* $G = \{1, 2\}$.

In particular we have

$$(6.1) \qquad \left| \bigcup_{k \geq 0} \mathbf{B}_k(g|w_{-k}^{-1}) \right| = \sum_{k \geq 0} |\mathbf{B}_k(g|w_{-k}^{-1})| = P(g|w_{-\infty}^{-1})$$

and

$$(6.2) \qquad \bigcup_{g \in G} \bigcup_{k \geq 0} \mathbf{B}_k(g|w_{-k}^{-1}) = [0, 1[,$$

where all the unions above are disjoint. For $k \geq 0$ let

$$\mathbf{B}_k(w_{-k}^{-1}) := \bigcup_{g \in G} \mathbf{B}_k(g|w_{-k}^{-1}).$$

By (2.5) and (4.9), we have

$$(6.3) \quad [0, a_k[\subset \bigcup_{\ell=0}^{k} \mathbf{B}_\ell(w_{-\ell}^{-1}) = \left[0, \sum_{g \in G} a_k(g|w_{-k}^{-1}) \right[\qquad \text{for all } w_{-\infty}^{-1} \in G^{-\mathbb{N}^*}.$$

As a consequence,

$$(6.4) \qquad [0, a_k[\cap \mathbf{B}_\ell(w_{-\ell}^{-1}) = \varnothing \qquad \text{for } \ell > k,$$

a fact that makes the definitions

$$(6.5) \qquad \mathbf{B}_{\ell,k}(w_{-1}, \ldots, w_{-k}) := [0, a_k[\cap \mathbf{B}_\ell(w_{-\ell}^{-1}),$$

$$(6.6) \qquad \mathbf{B}_{\ell,k}(g|w_{-1}, \ldots, w_{-k}) := [0, a_k[\cap \mathbf{B}_\ell(g|w_{-\ell}^{-1})$$

meaningful *for all* $\ell, k \geq 0$.

Items (ii) and (iii) [and the corresponding part in (v)] of Theorem 4.1 are immediate consequences of the following proposition. Similar to (3.5), define

$$\tau[n](\underline{u}) := \max\{s \leq n : u_j < a_{j-s}, \ j \in [s, n]\}.$$

PROPOSITION 6.1. *Let g_0 be an arbitrary point in G. Define the function $\Phi : [0, 1[^{\mathbb{Z}} \to G^{\mathbb{Z}}$, $\underline{u} \mapsto \underline{x} = \Phi(\underline{u})$, recursively from $j = \tau[n](\underline{u})$ to $j = n$ in the following fashion*: *write $\tau = \tau[n](\underline{u})$ and define*

$$x_\tau := \sum_{g \in G} g\, \mathbf{1}\{u_\tau \in \mathbf{B}_0(g)\},$$

$$x_{\tau+1} := \sum_{g \in G} g\, \mathbf{1}\{u_{\tau+1} \in \mathbf{B}_0(g) \cup \mathbf{B}_{1,1}(g|x_\tau)\},$$

(6.7)
$$\vdots$$

$$x_n := \sum_{g \in G} g\, \mathbf{1}\left\{u_n \in \bigcup_{0 \le j \le \tau} \mathbf{B}_{j,n-\tau}(g|x_{n-1}, \ldots, x_\tau)\right\}$$

if $\tau[n](\underline{u})$ is finite, and $x_n = g_0$ if $\tau[n](\underline{u}) = -\infty$. [Each sum in (6.7) reduces to a single term.] Then:

(i) *Φ is well defined and measurable.*

(ii) *For each $n \in \mathbb{Z}$, the component $\Phi(\underline{u})(n)$ depends only on the values of u_i on the interval $[\tau[n](\underline{u}), n]$ if $\tau[n](\underline{u}) > -\infty$, and on $(u_i : i \le n)$ otherwise.*

(iii) *If $\mathbb{P}(\tau[n](\underline{U}) > -\infty) = 1$, for all $n \in \mathbb{Z}$, then the law μ of $\Phi(\underline{U})$ is compatible with P and μ is stationary.*

PROOF.   Let

(6.8) $$A_n = \{\underline{u} \in [0, 1]^{\mathbb{Z}} : \tau[n](\underline{u}) > -\infty\}.$$

(i) On the set A_n, the consistency of the definition (6.7) follows from two facts: (1) By definition, $\tau[n] \le \tau[j]$ for $j \in [\tau[n], n]$, and (2), (6.3) shows that if the event $\{U_n < a_k\}$ holds, then we only need to look at $x_{n-1}, \ldots, x_{n-k}$ to obtain the value of x_n. These facts imply that for every $k \in [\tau[n], n]$, the value of x_k computed using (6.7) with k in place of n yields the same value as the one obtained as part of the recursive calculation (6.7) for x_n. The $\mathcal{F}(U_i : i \le n)$-measurability of x_n follows, then, from definition (6.7). As the sets A_n are $\mathcal{F}(U_i : i \le n)$-measurable [see (4.2)], the maps x_n remain measurable after gluing. We conclude that $\Phi(\underline{u}) = (x_n; n \in \mathbb{Z})$ is well defined and measurable.

(ii) Immediate from the definition (now known to be consistent).

(iii) When A_n is true, (6.4) implies that definition (6.7) amounts to

(6.9) $$x_n = \sum_{g \in G} g\, \mathbf{1}\left\{u_n \in \bigcup_{k \ge 0} \mathbf{B}_k(g|x_{n-k}^{n-1})\right\},$$

which implies

(6.10)
$$\mathbb{P}\big(\Phi(\underline{U})(n) = g \,\big|\, U_i : i < n;\ A_n\big)$$
$$= \mathbb{P}\left(U_n \in \bigcup_{k \ge 0} \mathbf{B}_k\big(g|[\Phi(\underline{U})]_{n-k}^{n-1}\big) \,\bigg|\, U_i : i < n;\ A_n\right).$$

Since each $\Phi(\underline{U})(j)$ is $\mathcal{F}(U_i : i \leq j)$-measurable, and each U_n is independent of $\mathcal{F}(U_i : i < n)$, if A_n has full measure, then (6.10) equals

$$(6.11) \quad \left| \bigcup_{k \geq 0} \mathbf{B}_k(g \mid [\Phi(\underline{U})]_{n-k}^{n-1}) \right| = \sum_{k \geq 0} \left| \mathbf{B}_k(g \mid [\Phi(\underline{U})]_{n-k}^{n-1}) \right| = P(g \mid [\Phi(\underline{U})]_{-\infty}^{n-1})$$

by (6.1). In other words,

$$(6.12) \qquad \mathbb{P}\big(\Phi(\underline{U})(n) = g \mid U_i : i < n\big) = P\big(g \mid \Phi(\underline{U})(i) : i < n\big).$$

As the right-hand side depends only on $(\Phi(\underline{U})(i), i < n)$, so does the left-hand side. Hence the law of $\Phi(\underline{U})(n)$ given $(\Phi(\underline{U})(i), i < n)$ itself is still given by P. Therefore the distribution μ of $\Phi(\underline{U})$ is compatible with P in the sense of (2.2). It is stationary by construction. $\square$

REMARK 6.1. The previous argument shows, in particular, that if $a_k^* \leq a_k$ for all $k \geq 0$, and $\tau^*[s, t](\underline{u})$ is finite (in which case $\tau[s, t] < \tau^*[s, t]$ is finite; see Remark 5.1), then $\Phi(\underline{u})(j) = \Phi^*(\underline{u})(j)$ for all $j \in [s, t]$. [Each $\Phi^*(\underline{u})(j)$ depends on a larger number $j - \tau^*[j]$ of preceding $U's$.] If condition (4.3) holds for the coefficients $\beta_m^* = \prod_{k=0}^m a_k^*$, both Φ^* and Φ have laws compatible with P.

PROOF OF COROLLARY 4.1. We represent the one-sided conditioned measure through a family of functions $\Phi(\cdot | \underline{w}, i) : [0, 1]^{\mathbb{Z}} \mapsto G^{\mathbb{Z}}$ corresponding to a history $\underline{w} \in G^{\mathbb{Z}}$ frozen for times smaller than $i \in \mathbb{Z}$. Writing, for shortness, $y_j = \Phi(\underline{U} | \underline{w}, i)(j)$, we set $y_j = w_j$ for $j < i$ and successively for $n = i, i+1, \ldots,$

$$(6.13) \quad y_n := \sum_{g \in G} g \mathbf{1} \left\{ U_n \in \bigcup_{k \geq 0} \mathbf{B}_k(g \mid y_{n-1}, \ldots, y_i, w_{i-1}, \ldots, w_{n-k}, \ldots) \right\}.$$

Let μ be any probability measure on $G^{\mathbb{Z}}$ compatible with P. From (2.3) and (6.1) we see that the law of $\Phi(\underline{U} | \underline{w}, i)$ is a regular version of μ given $\eta_{i-1} = \omega_{i-1}, \eta_{i-2} = \omega_{i-2}, \ldots$. That is,

$$(6.14) \qquad \mu\big(f \mid \eta(j) = w_j, \ j < i\big) = \mathbb{E}\big[f\big(\Phi(\underline{U} | w_{-\infty}^{-1}, i)\big)\big]$$

for any continuous f. We now follow the classical arguments.

PROOF OF (i). Let f be as stated. By (6.14),

$$(6.15) \quad \begin{aligned} &\mu\big(f \mid \eta(j) = w_j, \ j < i\big) - \mu\big(f \mid \eta(j) = v_j, \ j < i\big) \\ &\qquad = \mathbb{E}\Big[f\big(\Phi(\underline{U} | w_{-\infty}^{-1}, i)\big) - f\big(\Phi(\underline{U} | v_{-\infty}^{-1}, i)\big)\Big]. \end{aligned}$$

Since

$$(6.16) \qquad \mathbf{1}\big\{\Phi(\underline{U} | \underline{w}, i)(n) \neq \Phi(\underline{U} | \underline{v}, i)(n)\big\} \leq \mathbf{1}\big\{\tau[n] \leq i\big\},$$

the absolute value of both terms in (6.15) is bounded above by

$$(6.17) \qquad\qquad 2\|f\|_\infty \mathbb{P}(\tau[s,t] < i).$$

To conclude, we use the bound (4.6).

PROOF OF (ii). If μ and μ' are two measures on $G^\mathbb{Z}$ compatible with P,

$$
\begin{aligned}
|\mu f - \mu' f| &= \left| \int \mu(dw_{-\infty}^{-1})\mu(f \mid \eta(j) = w_j,\ j < i) \right. \\
&\qquad \left. - \int \mu'(d\underline{v})\mu'(f \mid \eta(j) = v_j,\ j < i) \right| \\
&\le \iint \mu(dw_{-\infty}^{-1})\mu'(d\underline{v}) \Big| \mu(f \mid \eta(j) = w_j,\ j < i) \\
&\qquad\qquad\qquad\qquad - \mu(f \mid \eta(j) = v_j,\ j < i) \Big|,
\end{aligned}
$$

(6.18)

which, by part (i), goes to zero as $i \to -\infty$. $\square$

PROOF OF THEOREM 4.1. We finish this section by showing how Theorem 4.1 follows from previous results.

Lemma 5.1 proves (i) and (iv) and the corresponding part of (v). The bounds mentioned at the end of (iv) are consequence of Proposition 5.1. Proposition 6.1 proves parts (ii) and (iii) and the corresponding part in (v). $\square$

7. Perfect simulation. In this section we prove Corollary 4.2. The construction of the function $\widetilde{\Phi}_{\tau,t}$ relies on an alternative construction of the stopping time $\tau[s,t]$. Assume $s \le t < \infty$ and define

$$(7.1) \qquad\qquad Z[s,t] := \max\{K_n - n + s : n \in [s,t]\} \ge 0,$$

where $K_n = K_n(\underline{U})$ is defined in (3.4). $Z[s,t]$ is the number of sites to the left of s we need to be able to construct Φ in the interval $[s,t]$. Let $Y_{-1} := t + 1$, $Y_0 := s$ and for $n \ge 1$, inductively,

$$(7.2) \qquad\qquad Y_n := Y_{n-1} - Z[Y_{n-1}, Y_{n-2} - 1].$$

Then it is easy to see that

$$(7.3) \qquad\qquad \tau[s,t] = \lim_{n\to\infty} Y_n = \max\{Y_n : Y_n = Y_{n+1}\} \qquad \text{a.s.,}$$

with the usual convention $\max \varnothing = -\infty$.

Construction of the perfect simulation function $\widetilde{\Phi}_{\tau,t}$.

1. Set $Y_{-1} = t + 1$, $Y_0 = s$ and iterate the following step until $Y_n = Y_{n-1}$:
   Generate $U_{Y_\ell}, \ldots, U_{Y_{\ell-1}-1}$. Use (3.4) to compute $K_{Y_\ell}, \ldots, K_{Y_{\ell-1}}$ and (7.1) and (7.2) to compute $Y_{\ell+1}$.
2. Let $\tau = Y_n$.
3. Iterate the following procedure from $k = \tau$ to $k = t$:
   Define x_k using (6.7).
4. Return $(x_j : j \in [s, t]) = (\widetilde{\Phi}_{\tau,t}(U_\tau, \ldots, U_t)(j) : j \in [s, t])$.

By definition of τ, $\widetilde{\Phi}_{\tau,t}(U_\tau, \ldots, U_t)(j) = \Phi(\underline{U})(j)$, $j \in [s, t]$.

REMARK 7.1. The algorithm can be applied for any choice of $(a_k^*, k \geq 0)$ satisfying (4.9) and, in addition,

$$\sum_m \prod_{k=0}^m a_k^* = \infty.$$

The smaller the a_k^*, the smaller the stopping times $\tau^*[s, t]$ of the resulting perfect simulation scheme. Also the return probabilities ρ_m^* increase if the a_k^* increase, worsening the bound (4.12) on the user-impatience bias.

8. Regeneration scheme. In this section we prove Corollary 4.3. The stationarity of **N** follows immediately from the construction. Let

(8.1) $$f(j) := \mathbb{P}(\mathbf{N}(-j) = 1 \mid \mathbf{N}(0) = 1)$$

for $j \in \mathbb{N}^*$. To see that **N** is a renewal process it is sufficient to show that

(8.2) $$\mathbb{P}(\mathbf{N}(s_\ell) = 1; \ell = 1, \ldots, n) = \beta \prod_{\ell=1}^{n-1} f(s_{\ell+1} - s_\ell)$$

for arbitrary integers $s_1 < \cdots < s_k$. [From Poincaré's inclusion–exclusion formula, a measure on $\{0, 1\}^{\mathbb{Z}}$ is characterized by its value on cylinder sets of the form $\{\zeta \in \{0, 1\}^{\mathbb{Z}} : \zeta(s) = 1, s \in S\}$ for all finite $S \subset \mathbb{Z}$. For $S = \{s_1, \ldots, s_k\}$, a renewal process must satisfy (8.2).] For $j \in \mathbb{Z}$, $j' \in \mathbb{Z} \cup \{\infty\}$, define

(8.3) $$H[j, j'] := \begin{cases} \{U_{j+\ell} < a_\ell, \ \ell = 0, \ldots, j' - j\}, & \text{if } j \leq j', \\ \text{"full event,"} & \text{if } j > j'. \end{cases}$$

With this notation,

(8.4) $$\mathbf{N}(j) = \mathbf{1}\{H[j, \infty]\}, \qquad j \in \mathbb{Z}$$

and

(8.5) $$\mathbb{P}(\mathbf{N}(s_\ell) = 1; \ell = 1, \ldots, n) = \mathbb{P}\left\{\bigcap_{\ell=1}^n H[s_\ell, \infty]\right\}.$$

From monotonicity we have for $j < j' < j'' \leq \infty$,

$$(8.6) \qquad H[j, j''] \cap H[j', j''] = H[j, j' - 1] \cap H[j', j''],$$

and then, with $s_{n+1} = \infty$ we see that (8.5) equals

$$(8.7) \qquad \prod_{i=1}^{n} \mathbb{P}\{H[s_\ell, s_{\ell+1} - 1]\}.$$

On the other hand,

$$(8.8) \qquad f(j) = \mathbb{P}(H[-j, \infty] \mid H[0, \infty]) = \mathbb{P}(H[-j, -1]).$$

Hence, (8.7) equals the right-hand side of (8.2) and we have proved that **N** is a renewal process.

On the other hand, by stationarity,

$$(8.9) \qquad \mathbb{P}(T_{\ell+1} - T_\ell \geq m) = \mathbb{P}(\tau[-1, \infty] < -m + 1 \mid \tau[0, \infty] = 0)$$

and, hence, by (5.6) and (5.3),

$$(8.10) \qquad \mathbb{P}(T_{\ell+1} - T_\ell \geq m) = \mathbb{P}(W_{-1}^{-m+1} = 0) = \rho_m,$$

proving (4.14).

The independence of the random vectors ξ_ℓ follows from the definition of T_ℓ and part (iii) of Theorem 4.1.

9. Applications.

9.1. *Binary autoregressions, continued.* In this subsection we continue the discussion of example (3.1). Recalling the notation of Section 3, we define

$$(9.1) \qquad C^+ = \max\left\{q'(x) : x \in \left[\theta_0 - \sum_{m>0} |\theta_m|, \theta_0 + \sum_{m>0} |\theta_m|\right]\right\},$$

$$(9.2) \qquad C^- = \min\left\{q'(x) : x \in \left[\theta_0 - \sum_{m>0} |\theta_m|, \theta_0 + \sum_{m>0} |\theta_m|\right]\right\}.$$

From the definition (4.8), a simple computation shows that

$$(9.3) \qquad \beta > 0 \iff \sum_k r_k < \infty \iff \sum_k k|\theta_k| < \infty$$

and also that for $|\theta_k| \sim Ck^{-2}$, condition (4.3) is satisfied for $C < (2C^+)^{-1}$, but not satisfied for $C > (2C^-)^{-1}$. Hence $\limsup_k k^2|\theta_k| < (2C^+)^{-1}$ is sufficient for being in the Harris regime (4.3), and $\limsup_k |\theta_k|k^2 \ln^2 k < \infty$ implies that $\beta > 0$.

Since a_k given by (3.3) has no simple expression in general, Remark 4.1 could be useful. Indeed,

$$(9.4) \qquad 1 - 2C^+ r_k \leq a_k \leq 1 - 2C^- r_k.$$

Under the extra condition that $\theta_k \neq 0$ for infinitely many k's, we have in fact $a_k \sim 1 - 2C^+ r_k$ as $k \to \infty$.

We can replace the coefficients a_k with the following lower bounds. We choose first some integer k_0 such that $2C^+ r_{k_0} < 1$ and define

$$(9.5) \qquad a_k^* = \begin{cases} a_k \wedge (1 - 2C^+ r_{k_0}) & \text{for } k < k_0, \\ 1 - 2C^+ r_k & \text{for } k \geq k_0. \end{cases}$$

We can use the modification of our algorithm at the end of Section 7 with these coefficients a_k^*. Note that we only need to compute at most k_0 different a_k's.

We now discuss two well-studied choices for q.

SIGMOÏD CASE. In addition we assume here that q is concave on $\mathbb{R}^+$ with $q(x) + q(-x) = 1$, $x \in \mathbb{R}$. One natural choice is

$$(9.6) \qquad q(x) = \frac{\exp x}{2 \cosh x} = \frac{1}{2(1 + \exp(-2x))},$$

that is, the so-called logistic function and logit model [Guyon (1995), Example 2.2.4], where the explicative variables are the values of the process in the past. For a general sigmoïd q, the supremum in (3.3) is achieved for w_{-k}^{-1} minimizing $|\theta_0 + \sum_{1 \leq m \leq k} \theta_m w_{-m}|$.

LINEAR CASE. We take now $q(x) = (1 + x)/2$, and necessarily $|\theta_0| + \sum_{m>0} |\theta_m| < 1$. As we will see, linearity makes the model (3.1) somehow trivial, but also instructive. Writing $\mathcal{B}(p)$ for the Bernoulli distribution with parameter p and δ the Dirac measure, we note that the one-sided conditional law (3.1) is given by a convex combination,

$$P(\cdot | w_{-\infty}^{-1}) = \mathcal{B}\left(\frac{1 + \theta_0 + \sum_{k \geq 1} \theta_k w_{-k}}{2} \right)$$

$$(9.7) \qquad = \mathcal{B}\left(\frac{1 + \theta_0 - r_0}{2} + \sum_{k \geq 1} |\theta_k| \frac{\text{sign}(\theta_k) w_{-k} + 1}{2} \right)$$

$$= (1 - r_0) \mathcal{B}\big((1 + \theta_0 - r_0)/2(1 - r_0) \big) + \sum_{k \geq 1} |\theta_k| \delta_{\text{sign}(\theta_k) w_{-k}},$$

since $\mathcal{B}(\lambda p + (1 - \lambda) p') = \lambda \mathcal{B}(p) + (1 - \lambda) \mathcal{B}(p')$ for $\lambda \in [0, 1]$. In this example we have $a_k(w_{-k}^{-1}) = 1 - r_k = a_k$ independent of w, and for $k \geq 1$,

$$(9.8) \ b_k(\pm 1 | w_{-k}^{-1}) = (\theta_k w_{-k} + |\theta_k|)/2 = \begin{cases} |\theta_k|, & \text{if } w_{-k} = \text{sign}(\theta_k), \\ 0, & \text{otherwise,} \end{cases} \in \{0, |\theta_k|\}.$$

Hence one of the two intervals $\mathbf{B}_k(\pm 1 | w_{-k}^{-1})$ is empty, while the other one $\mathbf{B}_k(\text{sign}(\theta_k) w_{-k} | w_{-k}^{-1})$ has length $|\theta_k|$. This is in accordance with formula (9.7).

In other respects, the decomposition (9.7) can be directly interpreted in terms of simulation: the value of the process at time $i = 0$ is chosen according to a "new" coin tossing with probability $1 - r_0$, and set to the value $\text{sign}(\theta_k)w_{-k}$ with probability $|\theta_k|$ ($k = 1, 2, \ldots$). In the latter case the value of w_{-k} is needed, and will be constructed using (9.7) again, etc. Clearly this recursive construction will eventually stop if and only if $|\theta_0| + \sum_{m>0} |\theta_m| < 1$. However, our construction in this paper requires the extra condition (4.3), which, loosely speaking, amounts to $\limsup_k k^2 |\theta_k| < 1$. The reason for stronger assumptions is that, in order to cover general processes, we need in our general construction to check *all intermediate times* between 0 and $\tau[0]$, though in the special case of P given by (9.7), it is not necessary to construct all of them following the above lines.

9.2. *Markov chains defined by D-ary expansions.* These are processes having the unit interval as "alphabet", $I = [0, 1]$, and defined through another, auxiliary, process with a finite alphabet. Formally, a family of maps is established between sequences of a finite alphabet $G = \{0, 1, \ldots, D - 1\}$ and real numbers in I via D-ary expansions. For each $n \in \mathbb{Z}$,

$$X_n : G^{\mathbb{Z}} \to I,$$

(9.9)
$$\big(\eta(i) : i \in \mathbb{Z}\big) \mapsto x_n = \sum_{j=1}^{\infty} \eta(n - j)/D^j.$$

This map induces a natural map from probability kernels $P : G \times G^{-\mathbb{N}^*} \mapsto [0, 1]$ to probability kernels $F : I \times I \mapsto [0, 1]$: For each $x \in I$, given an $w_{-\infty}^{-1} \in G^{-\mathbb{N}^*}$ with $x = X_0(w_{-\infty}^{-1})$,

(9.10)
$$F\left(X_1 = \frac{g + x}{D} \,\Big|\, X_0 = x\right) = P(g|\underline{w}).$$

Interest focuses on the existence and properties of measures on the Borelians of $I^{\mathbb{Z}}$ compatible with such a (Markov) kernel F.

Maps (9.9) and (9.10) have been introduced by Borel (1909) for i.i.d. $\eta(i)$. The general case in which the $\eta(i)$ form a chain with long memory is the object of Harris' (1955) seminal paper. They are the prototype of the random systems with complete connections mentioned in the Introduction. Harris determines conditions for the existence and uniqueness of these processes, through the study of long-memory chains: If the finite-alphabet chain satisfies a condition similar to (but weaker than) (4.3), there is a unique process $\mathcal{M}$ on $I^{\mathbb{Z}}$ compatible with F. This process $\mathcal{M}$ is of course a (stationary) Markov chain with transition probability kernel F. Harris shows that its marginal distribution is continuous, except in the degenerate case with constant $\eta(i)$'s where it is concentrated on one of the points $0, 1/(D-1), 2/(D-1), \ldots, 1$. Furthermore, if the process is mixing and not degenerate, this marginal is purely singular whenever the variables

$\eta(i)$ are not independent uniformly distributed (in which case the marginal is uniform).

Our approach yields, in a straightforward way, a perfect simulation scheme for the measures $\mathcal{M}$ obtained in this fashion, if the auxiliary process $\eta(i)$ satisfies condition (4.3). Indeed, the map

$$(9.11) \qquad \begin{aligned} &X : G^{-\mathbb{N}^*} \to I, \\ &w_{-\infty}^{-1} \mapsto x = X_0(w_{-\infty}^{-1}) \end{aligned}$$

can be made bijective by fixing rules to decide between sequences which are eventually 0 and those that are eventually $D - 1$. In turns, this map induces a bijection between the sigma algebra $\mathcal{S}_\ell$ formed by unions of intervals with endpoints in multiples of $D^{-\ell}$, and the subsets of $G^{[-\ell,-1]}$. We conclude that, if μ and $\mathcal{M}$ are the processes compatible with the kernels P and F related as in (9.10), then the restriction of $\mathcal{M}$ to $\mathcal{S}_\ell$ can be perfectly simulated by mapping, via X, the perfect samples of the measure μ on the window $[-\ell, -1]$, obtained by the algorithm of Section 7. We point out that the union of the families of $\mathcal{M}_\ell$ uniquely determines the measure $\mathcal{M}$ (it forms a so-called π system).

In conclusion, for Markov chains in general state-space with transition kernel $P(x, \cdot)$ satisfying the Doob's condition $P(x, \cdot) \geq \beta\varphi(\cdot)$, for all state x, some $\beta > 0$ and a measure φ on the state-space, the forward coupling is well known. The corresponding coupling-from-the-past algorithm is illustrated in Example 2 of Foss and Tweedy (1998) and in Corcoran and Tweedie (2001). Notice, however, that the mere existence of a minorization measure is not sufficient to construct the couplings: One needs to explicitly know φ and β. In this section we have discussed an example of a Markov chain with state-space $[0, 1]$ that can be transcribed as a chain with complete connections and state-space $\{0, 1, \ldots, D-1\}$. Perfect simulating the latter provides a perfect simulation for the former. Instead of the exibition of a minorization measure for the Markov chain, our method requires the knowledge of $a_k(g|w_{-\infty}^{-1})$ and lower bounds of a_k for the related infinite-memory chain.

REFERENCES

ATHREYA, K. and NEY, P. (1978). A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245** 493–501.

BARNSLEY, M., DEMKO, S., ELTON, J. and GERINOMO, J. (1988). Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. *Ann. Inst. H. Poincaré, Probab. Statist.* **24** 367–394. [Erratum (1989) **25** 589–590.]

BERBEE, H. (1987). Chains with complete connections: Uniqueness and Markov representation. *Probab. Theory Related Fields* **76** 243–253.

BLACKWELL, D. (1957). The entropy of functions on finite-state Markov chains. In *Trans. First Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* 13–20. Czechoslovak Akad. Sci., Prague.

BOREL, E. (1909). Sur les probabliités dénombrables et leurs applications arithmetiques. *Rend. Circ. Mat. Palermo* **26** 247–271.

BRESSAUD, X., FERNÁNDEZ, R. and GALVES, A. (1999). Decay of correlations for non Hölderian dynamics. A coupling approach. *Electron. J. Probab.* **4**.

CORCORAN, J. N. and TWEEDIE, R. L. (2001). Perfect sampling of Harris recurrent Markov chains. *Ann. Appl. Probab.* **11** 438–451.

DACUNHA-CASTELLE, D., DUFLO, M. and GENON-CATALOT, V. (1983). *Exercices de Probabilités et Statistiques: 2-Problèmes à Temps Mobile*. Masson, Paris.

DOEBLIN, W. (1938). Sur deux problémes de M. Kolmogoroff concernant les chaînes dénombrables. *Bull. Soc. Math. France* **66** 210–220.

DOEBLIN, W. (1940). Remarques sur la théorie métrique des fractions continues. *Compositio Math.* **7** 353–371.

DOEBLIN, W. and FORTET, R. (1937). Sur les chaînes à liaisons complétes. *Bull. Soc. Math. France* **65** 132–148.

ELTON, J. and PICCIONI, M. (1992). Iterated function systems arising from recursive estimation problems. *Probab. Theory Related Fields* **91** 107–114.

FERRARI, P. A., MAASS, A., MARTÍNEZ, S. and NEY, P. (2000). Cesàro mean distribution of group automata starting from measures with summable decay. *Ergodic Theory Dynam. Systems* **20** 1657–1670.

FILL, J. A. (1998). An interruptible algorithm for perfect sampling via Markov chains. *Ann. Appl. Probab.* **8** 131–162.

FOSS, S. G. and TWEEDIE, R. L. (1998). Perfect simulation and backward coupling. *Comm. Statist. Stochastic Models* **14** 187–203.

GUYON, X. (1995). *Random Fields on a Network*. Springer, New York.

HARRIS, T. E. (1955). On chains of infinite order. *Pacific J. Math.* **5** 707–724.

HERKENRATH, U. and THEODORESCU, R. (1978). General control systems. *Inform. Sci.* **14** 57–73.

IOSIFESCU, M. (1978). Recent advances in the metric theory of continued fractions. In *Trans. Eighth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* **A** 27–40. Reidel, Dordrecht.

IOSIFESCU, M. (1992). A coupling method in the theory of dependence with complete connections according to Doeblin. *Rev. Roumaine Math. Pures Appl.* **37** 59–65.

IOSIFESCU, M. and THEODORESCU, R. (1969). *Random Processes and Learning*. Springer, Berlin.

IOSIFESCU, M. and GRIGORESCU, S. (1990). *Dependence with Complete Connections and Its Applications*. Cambridge Univ. Press.

KAIJSER, T. (1975). A limit theorem for partially observed Markov chains. *Ann. Probab.* **3** 677–696.

KAIJSER, T. (1981). On a new contraction condition for random systems with complete connections. *Rev. Roumaine Math. Pures Appl.* **26** 1075–1117.

KAIJSER, T. (1994). On a theorem of Karlin. *Acta Appl. Math.* **34** 51–69.

KEANE, M. (1971). Sur les mesures invariants d'un recouvrement régulier. *C. R. Acad. Sci. Paris* A585–A587.

KEANE, M. (1976). Strongly mixing *g*-measures. *Invent. Math.* **16** 309–324.

KOZLOV, O. K. (1974). Gibbs description of a system of random variables. *Problems Inform. Transmission* **10** 258–265.

LALLEY, S. P. (1986). Regeneration representation for one-dimensional Gibbs states. *Ann. Probab.* **14** 1262–1271.

LALLEY, S. P. (2000). Regeneration in one-dimensional Gibbs states and chains with complete connections. *Resenhas* **4** 249–281.

LEDRAPPIER, F. (1974). Principe variationnel et systèmes dynamiques symboliques. *Z. Wahrsch. Verw. Gebiete* **30** 185–202.

MCCULLAGH, P. and NELDER, J. A. (1989). *Generalized Linear Models*, 2nd ed. Chapman-Hall, London.

NORMAN, F. (1972). *Markov Processes and Learning Models*. Academic Press, New York.

NORMAN, F. (1974). Markovian learning processes. *SIAM Rev.* **16** 143–162.

NORMAN, F. (1975). An ergodic theorem for evolution in a random environment. *J. Appl. Probab.* **12** 661–672.

NUMMELIN, E. (1978). A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **43** 309–318.

ONICESCU, O. and MIHOC, G. (1935a). Sur les chaînes statistiques. *C. R. Acad. Sci. Paris* **200** 511–512.

ONICESCU, O. and MIHOC, G. (1935b). Sur les chaînes de variables statistiques. *Bull. Sci. Math.* **59** 174–192.

PROPP, J. G. and WILSON, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* **9** 223–252.

PRUSCHA, H. and THEODORESCU, R. (1977). Functions of event variables of a random system with complete connections. *J. Multivariate Anal.* **7** 336–362.

WILSON, D. B. (1998). Annotated bibliography of perfectly random sampling with Markov chains. In *Microsurveys in Discrete Probability* (D. Aldous and J. Propp, eds.) **41** 209–220. Amer. Math. Soc., Providence, RI. (Updated versions available at http://dimacs.rutgers. edu/~dbwilson/exact.)

F. COMETS
MATHÉMATIQUES, CASE 7012
UNIVERSITÉ PARIS 7 – DENIS DIDEROT
2 PLACE JUSSIEU
75251 PARIS CEDEX 05
FRANCE
E-MAIL: comets@math.jussieu.fr

R. FERNÁNDEZ
CNRS-UPRES-A 6085 MATHÉMATIQUES
UNIVERSITÉ DE ROUEN
SITE COLBERT
F 76821 MONT SAINT AIGNAN
FRANCE
E-MAIL: Roberto.Fernandez@univ-rouen.fr

P. A. FERRARI
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE DE SÃO PAULO
CAIXA POSTAL 66281
05315-970 SÃO PAULO
BRAZIL
E-MAIL: pablo@ime.usp.br