A general framework for perfect simulation of long memory processes

Emilio De Santis
Dipartimento di Matematica
Sapienza Università di Roma
desantis@mat.uniroma1.it

Mauro Piccioni
Dipartimento di Matematica
Sapienza Università di Roma
piccioni@mat.uniroma1.it

April 2, 2010

Abstract

In this paper a general approach for the perfect simulation of a stationary process with at most countable state space is outlined. The process is specified through a kernel, prescribing the probability of each state conditional to the whole past history. We follow the seminal paper [CFF02], where sufficient conditions for the construction of a certain perfect simulation algorithm have been given. We generalize this approach by defining backward coalescence times for these kind of processes; this allows us to construct perfect simulation algorithms under weaker conditions. Backward coalescence times are constructed in the following ways: (i) by taking into account some a priori knowledge about the histories that occur; (ii) by merging the algorithm in [CFF02] with the classical CFTP algorithm [PW96].

Keywords: Perfect simulation, Coupling, Renewal processes.
AMS classification: 60G10, 60J10.

1 Introduction

Perfect simulation algorithms for stochastic processes have been developed mostly for Markov chains, starting from the original CFTP algorithm presented in the founding paper by Propp and Wilson [PW96]. Later on, Foss and Tweedie [FT98] recognized the fundamental role of the so-called stochastic recursive sequences for perfect simulation. Murdoch and Green
constructed a stochastic recursive sequence for perfect simulation when the transition kernel satisfies a minorization condition, called the gamma-coupler. Among the applications of perfect simulation in recent years, we cite stochastic geometry ([FFG02], [Møl01]) and random fields ([HS00], [DSP08]).

The fact that the main idea underlying stochastic recursive sequences, and the gamma-coupler in particular, works beyond the markovian case, is shown by the extension, due to Comets et al [CFF02], to processes with infinite memory. The aim of the present work is to present some generalizations of their results.

In this paper we consider stochastic processes defined on $\mathbb{Z}$ with values in an alphabet $G$, which is finite or countable: thus realizations of these processes are two-sided infinite words. The law of the process is obtained through a transition kernel prescribing the probability that each letter of the alphabet occurs in any given position of the word, conditional to the whole history preceding it.

For notational convenience we use the following convention concerning sequences with values in $G$: whenever $m \leq n$ are elements of $\mathbb{Z}$ we define the word

$$s^m_n = (s_n, \ldots, s_m) \in G^{m-n+1}.$$  

With an obvious extension we also allow $m = -\infty$ and $n = +\infty$. For brevity of notation we write $s := s^+\infty$. For $m = -\infty$ and $n$ finite the word will be called a *history*. Histories are elements of $G^{-N^*}$, where $N^*$ is the set of positive integers. We can concatenate any word $s^m_n$ with a history $w^{-1}_{-\infty}$, obtaining another history

$$(s^m_n, w^{-1}_{-\infty}) = (s_n, \ldots, s_m, w_{m-1}, w_{m-2}, \ldots).$$

The same notational conventions are used for sequences with values in the interval $[0, 1)$.

The set of histories $G^{-N^*}$ is equipped with the ultrametric distance

$$\delta(s^{-1}_{-\infty}, t^{-1}_{-\infty}) = 2^{-\inf\{n: s_n \neq t_n\}}, \quad s^{-1}_{-\infty}, t^{-1}_{-\infty} \in G^{-N^*}.$$  

The corresponding Borel $\sigma$-algebra coincides with the product $\sigma$-algebra, which is generated by all cylinder sets.

Let us consider a kernel $p : G \times G^{-N^*} \to [0, 1]$, which will be denoted by $p(g|w^{-1}_{-\infty})$. This means that, for any $g \in G$, $p(g|\cdot)$ is a measurable function in $G^{-N^*}$ such that $\sum_{g\in G} p(g|w^{-1}_{-\infty}) = 1$.

We say that a process $\{X_n, n \in \mathbb{Z}\}$ is *compatible* with the kernel $p$, if for any $m \in \mathbb{Z}$ and $g \in G$

$$P(X_m = g|X_{m-i}, i \in N^*) = p(g|X^{-1}_{-\infty}), \quad \text{a.s.} \quad (1)$$

For any $n \in \mathbb{N}_+$, the "one-dimensional" kernel $p$ induces an $(n+1)$-dimensional kernel
$p^{(n+1)} : G^{n+1} \times G^{-N_+} \to [0, 1]$ defined by

$$p^{(n+1)}(g_n, \ldots, g_0|w_{-\infty}) = \prod_{k=0}^{n} p(g_k|g_{k-1}, \ldots, g_0, w_{-\infty}), \quad g_k \in G, \quad k = 0, 1, \ldots, n, \quad w_{-\infty}^{-1} \in G^{-N_+}. \tag{2}$$

If the process \( \{X_n, n \in \mathbb{Z}\} \) is compatible with \( p \), then for any \( m \in \mathbb{Z}, n \in \mathbb{N} \) and any choice of \( g_k \in G, k = 0, \ldots, n \)

$$P(X_{m+k} = g_k, k = 0, \ldots, n|X_{m-i}, i \in \mathbb{N}^*) = \prod_{k=0}^{n} p(g_k|g_{k-1}, \ldots, g_0, X_{m-i}^{-1}) \quad \text{a.s.} \tag{3}$$

Starting from a kernel \( p \), perfect simulation is aimed to construct algorithms for sampling a compatible process \( \{X_n, n \in \mathbb{Z}\} \), giving at the same time sufficient conditions for its uniqueness.

Processes of this type are known in the literature as random systems with complete connections: for the foundations of their theory see [IG90], whereas for a full account about the literature on these processes we refer to the introduction in [CFF02]. In particular, previously known uniqueness conditions were given in [Ber87] and [Lal00].

We conclude this introduction by giving a plan of the paper. In Section 2 we introduce general coupling functions in the context of processes with infinite memory and define a general backward coalescence time. Its existence allows to deduce perfect simulation algorithms for the unique stationary process compatible with the kernel. At the end of the section the "maximal" coupling function is described, as introduced in [CFF02].

From the nature of the backward coalescence time used in [CFF02] we abstract the notion of information depth, to which Section 3 is devoted. This is a stopping time, associated to each instant and adapted to the past values of the random sources feeding the coupling function, which bounds the amount of information needed on the previous states in order to compute the state at that instant. From an information depth we give a canonical way to construct a backward coalescence time, provided it can be shown that it is a.s. finite. As in [CFF02], under slightly stronger condition, there is also a regeneration structure, which can be exploited to simulate the process for all positive times and not just on a finite window.

In Section 4 we modify the information depth used in [CFF02] in order to deal with examples like alternating renewal process and more general processes with a random change of regime in the memory of the past, for which the algorithm in [CFF02] is not successful. The main ingredient for proving that our modification works is the knowledge of the histories that could occur, once the sources of randomness are produced backward in time.

Finally, in Section 5 we construct directly a backward coalescence time, when the information depth in [CFF02] is always strictly positive, denying the possibility of building from it a backward coalescence time. In general, the construction of this modified backward coalescence times needs the modification of the coupling function as well. The construction
also requires a positive probability of coalescence in what we call the markovian regime: under suitable conditions, a perfect simulation algorithm is constructed by combining the algorithm in [CFF02] with the classical CFTP in [PW96]. A class of examples in which these conditions hold is finally discussed.

2 Coupling functions, backward coalescence times and perfect simulation

In this section we discuss some general issues involved in the design of a perfect simulation algorithm for a process compatible with a kernel $p : G \times G^{-N^*} \rightarrow [0, 1]$ of the form described above. The first concept to be introduced is that of coupling function. Despite the fact that in all the examples presented throughout the paper the coupling function is almost always the same, borrowed from [CFF02], we believe that it is useful to give an abstract definition. In particular we choose to make explicit the dependence of the backward coalescence time, which is defined afterwards, on the coupling function. In fact, several backward coalescence times will be discussed throughout the paper, different from the one used in [CFF02], for the same coupling function.

We first give the definition of admissible history, which is related to the zeros of the kernel function. This concept will be useful in the definitions and the results that follow.

We start with the definition of a forbidden word of the alphabet $G$, recursively on the length. A letter $g$ of the alphabet $G$ is forbidden if $p(g|w_{-\infty}^{-1}) = 0$, for any $w_{-\infty}^{-1} \in G^{-N^*}$. A word $s_{-n}^0 = (s_0, \ldots, s_{-n+1}, s_{-n})$ of length $n + 1$ is forbidden if either $s_{-n}^{-1}$ is forbidden or

$$p(s_0|w_{-\infty}^{-1}) = 0, \quad \forall w_{-\infty}^{-1} \in G^{-N^*} : w_{-n}^{-1} = s_{-n}^{-1}.$$ 

For any $n \in \mathbb{N}_+$, we call $\mathcal{H}_n$ the set of words that are not forbidden, of length $n$.

Next define the set of admissible histories

$$\mathcal{H} = \{ w_{-\infty}^{-1} \in G^{-N^+} : w_{-n}^{-1} \in \mathcal{H}_n, n \in \mathbb{N}_+ \}.$$ 

Since the set of histories $w_{-\infty}^{-1} \in G^{-N^*}$ such that $w_{-n}^{-1} = s_{-n}^{-1}$ for any fixed $s_{-n}^{-1} \in G^{n+1}$ is closed, the set $\mathcal{H}$, being an intersection of sets of this form, is closed.

Now we prove that a process $X$, which is compatible with $p$, is such that $X_{0-\infty}$ belongs to $\mathcal{H}$ with probability 1. It is readily seen that it is enough to prove that if $s_{-n}^0$ is a forbidden word for $p$, then

$$P(X_0 = s_0, \ldots, X_{-n} = s_{-n}) = 0.$$  \hspace{1cm} (4)$$

In order to prove this, notice that by the definition of forbidden word and (2) it is

$$p^{(n+1)}(s_0, \ldots, s_{-n}|w_{-\infty}^{-n-1}) = 0,$$
for any $w_{-∞}^{-n-1} \in G^{-N^*}$. By integrating over $w_{-∞}^{-n-1}$ with respect to the law of the process we conclude that (1) holds. With exactly the same argument it is proved that $X^m_{-∞}$ belongs to $H$ w.p. 1 as well, for any $m \in Z$.

In principle, the construction of the set of admissible histories can be iterated, replacing in the above definitions $G^{-N^*}$ with $H$. In this way the set of forbidden words could be enlarged, and thus the set of admissible histories could be reduced, and so on. We choose not to pursue this kind of generalization, since the previous definition is adequate for the examples which will be presented during the paper.

A coupling function $f$ for the kernel $p$ is a function $f : [0,1) \times H \rightarrow G$ such that for any $w_{-∞}^{-1} \in H$ and any $g \in G$ the set \( \{ u \in [0,1) : f(u|w_{-∞}^{-1}) = g \} \) is a disjoint union of intervals $[c_i(g|w_{-∞}^{-1}), d_i(g|w_{-∞}^{-1}))$, $i \in N_+$, of total length $p(g|w_{-∞}^{-1})$. For practical simulation purposes we also assume that $c_i(g|w_{-∞}^{-1})$ and $d_i(g|w_{-∞}^{-1})$ can be computed by looking at a finite portion of the history $w_{-∞}^{-1}$. This implies that $f$ is jointly measurable and for any $u \in [0,1)$ the function $f(u|\cdot)$ is continuous in $H$.

Since an interval $[c,d)$ is either empty or it has positive length, the set of admissible histories is invariant under the coupling function, in the sense that for any $u \in [0,1)$

$$w_{-∞}^{-1} \in H \Rightarrow (f(u|w_{-∞}^{-1}), w_{-∞}^{-1}) \in H.$$ 

As a consequence of the definition, if $f$ is a coupling function for $p$ and $U$ is a random variable uniformly distributed in $[0,1)$, then

$$P(f(U|w_{-∞}^{-1}) = g) = p(g|w_{-∞}^{-1}), \text{ for each } g \in G, \ w_{-∞}^{-1} \in H. \quad (5)$$

Starting from $f^{(1)} = f$, we define recursively $f^{(n)} : [0,1)^n \times H \rightarrow G$ for any $n \in N$, in the following natural way

$$f^{(n+1)}(u_n, \ldots, u_0|w_{-∞}^{-1}) := f \left( u_n | f^{(n)} \left( u_{n-1}, \ldots, u_0 | w_{-∞}^{-1} \right) \right), \ldots, f^{(1)}(u_0 | w_{-∞}^{-1}), w_{-∞}^{-1}. \quad (6)$$

Thus, whenever $U_k, k = 0, \ldots, n$ are i.i.d. random variables with the uniform distribution in $[0,1)$, the random vector $(f^{(k+1)}(U_k, \ldots, U_0|w_{-∞}^{-1}), k = 0, \ldots, n)$ has the law $p^{(n+1)}$ given in (2), for any $w_{-∞}^{-1} \in H$. This means that through the iterations of the coupling function it is possible to define, in the same probability space, a family of processes evolving in forward time according to the given kernel $p$, indexed by all admissible histories $w_{-∞}^{-1} \in H$.

For the implementation of a perfect simulation algorithm we require a coupling function $f$ to admit a backward coalescence time, which we are going to define.

Consider a sequence $U = \{U_i, i \in Z\}$ of i.i.d. random variables, with uniform distribution in the interval $[0,1)$, used as the source of randomness for the construction of the processes of interest. For any $m, n \in Z$ with $m \leq n$ we define the $σ$-algebra $\mathcal{F}_m^n = σ(U_i, i = m, \ldots, n)$. For simplicity of notation we specify an arbitrary reference admissible history $g_{-∞}^{-1} \in H$. We say that a measurable function $τ_0(U^0_{-∞})$ with non positive values is a backward coalescence time if it has the properties:
H1. $\tau_0$ is an a.s. finite stopping time w.r.t. the filtration $\{\mathcal{F}_n^0 : n \in \mathbb{N}\}$, i.e. $\{\tau_0 = -l\} \in \mathcal{F}_l^0$ for any $l \in \mathbb{N}$;

H2. if $\tau_0 = -l$, then for any $w_{-\infty}^{-(l+1)} \in \mathcal{H}$

$$f^{(l+1)}(U_0, U_{-1}, ..., U_{-l} | w_{-\infty}^{-(l+1)}) = f^{(l+1)}(U_0, U_{-1}, ..., U_{-l} | g_{-\infty}^{-(l+1)}).$$

The meaning of this definition is that by pushing back the initial time until $\tau_0$, which is computable by simulating, backward in time, the sequence $\{U_0, U_{-1}, \ldots\}$, the dependence of the value of the coupling function at time 0 on the history prior to time $\tau_0$ vanishes: only the dependence on $\{U_0, U_{-1}, \ldots, U_{\tau_0}\}$ remains. It is readily seen that property H2 remains true for $\tau_0 > -l$.

Likewise we can repeat the same construction for any $n \in \mathbb{Z}$, defining

$$\tau_n(U^n_{-\infty}) := n + \tau_0(U^n_{-\infty}), \forall n \in \mathbb{Z}. \quad (8)$$

If $\tau_0$ is a.s. finite, $\tau_n$ is finite as well, by translation invariance. For further use, for $m \leq n$, we also define

$$\tau[m, n] = \inf\{\tau_m, \tau_{m+1}, \ldots, \tau_{n-1}, \tau_n\}, \quad m \leq n. \quad (9)$$

If $\tau_0$ is a backward coalescence time we define the process $X = \{X_n, n \in \mathbb{Z}\}$ as

$$X_n = \sum_{l \in \mathbb{N}} f^{(l+1)}(U_n, U_{n-1}, ..., U_{n-l} | g_{-\infty}^{n-l-1}) 1_{\{\tau_n = n-l\}}(U^n_{-\infty}). \quad (10)$$

Notice that the definition does not depend on the choice of the reference history $g_{-\infty}^{-1} \in \mathcal{H}$.

**Proposition 1.** If $\tau_0$ is a backward coalescence time, then the process $X$ is stationary and it is the unique process compatible with the kernel $p$.

**Proof.** The stationarity of $X$ is guaranteed by construction. Let us proceed to prove that it is compatible with $p$. By stationarity it is enough to prove (I) for $m = 0$. Given the realization $U^0_{-\infty}$ define the non empty random subsets of $\mathcal{H}$

$$I_{n+1}(U_0, U_{-1}, \ldots, U_{-n}) = \{(f^{(n+1)}(U_0, \ldots, U_{-n} | w_{-\infty}^{-(n+1)}), \ldots, f^{(1)}(U_{-n} | w_{-\infty}^{-(n+1)}), w_{-\infty}^{-(n+1)} : w_{-\infty}^{-(n+1)} \in \mathcal{H}\},$$

for $n \in \mathbb{N}$, made of admissible histories in $\mathcal{H}$, obtained by varying in all possible ways the initial history prior to time $-n$, and then applying the coupling function with the fixed values $U_{-n}, \ldots, U_0$, until time zero.
The sequence $I_{n+1}(U_0, U_{-1}, \ldots, U_{-n})$ is non increasing in $n \in \mathbb{N}$. Moreover each element in $I_{-\tau[-n,0]+1}(U_0, U_{-1}, \ldots, U_{-n,0})$ has ultrametric distance from $X^0_{-\infty}$ which does not exceed $2^{-n}$. Therefore $X^0_{-\infty}$ belongs to the closure of

$$I_1(U_0) = \{(f(U_0|w^{-1}_{-\infty}), w^{-1}_{-\infty}) : w^{-1}_{-\infty} \in \mathcal{H}\}.$$ 

But the continuity of $f(u|\cdot)$ implies that $I_1(U_0)$ is closed, hence $X^0_{-\infty} \in I_1(U_0)$, meaning that $X^{-1}_{-\infty} \in \mathcal{H}$ and

$$X_0 = f(U_0|X^{-1}_{-\infty}).$$

Since $U_0$ is independent of $\mathcal{F}_{-\infty}$ and $X_{-n}$ is measurable w.r.t. this $\sigma$-algebra, for any $n \in \mathbb{N}_+$, this implies that $X$ satisfies (11) with $m = 0$. The proof of uniqueness is essentially the same as in [CFF02] p. 935. It is a consequence of the fact that the tail probability $P(\tau [0,n] \leq -i)$, gives an upper bound on the variation distance between two distributions of the form (2) indexed by any two initial histories in $\mathcal{H}$ which differ only before time $-i$. By a.s. finiteness of $\tau [0,n]$, this tail probability goes to zero as $i \to \infty$. \[\square\]

The construction of the process $X$ yields a perfect simulation algorithm on a finite window $[m,n]$, obtained first by a backward inspection of the sequence $(U_{n-i}, i \in \mathbb{N})$ in order to locate the stopping time $\tau [m,n]$ defined in (9), and then by a recursive computation of the coupling function started from the reference initial history $g_{\tau [m,n]-1}$. In general we cannot say that all the intermediate values of $X$ prior to time $m$ are identified during this computation, but this will happen for the kind of backward coalescence times considered in the next section.

Mimicking the proof of Proposition [1] we can construct the process $X$ also when the set of admissible histories $\mathcal{H}$ is replaced by a possibly smaller subset $\mathcal{H}'$ in the definition of the coupling function $f$ and the backward coalescence time $\tau_0$.

**Proposition 2.** Suppose that:

1. $\mathcal{H}'$ is invariant under the coupling function and under the cut operator sending $w^{-1}_{-\infty}$ into $w^{-2}_{-\infty}$;
2. for any $u \in [0,1)$, the function $f(u|\cdot)$ is continuous in $\mathcal{H}'$;
3. $X \in \mathcal{H}'$ with probability 1;
4. any $P$ compatible with the kernel $p$ gives probability 1 to $\mathcal{H}'$.

Then, if $\tau_0$ satisfies the assumptions H1 and H2 (with $\mathcal{H}$ replaced by $\mathcal{H}'$), the process $X$ is stationary and it is the unique process compatible with the kernel $p$. 

7
Finally we present the construction of the "maximal" coupling function introduced in [CFF02], modified by taking into account only the trajectories in the admissible set of histories \( \mathcal{H} \). In order to present this coupling function some relevant quantities have to be defined. First define \( a_k : G \times \mathcal{H} \to [0,1] \) as
\[
a_0(g) = \inf \{ p(g|z_{-\infty}^{-1}) : z_{-\infty}^{-1} \in \mathcal{H} \}, \quad g \in G,
\]
\[
a_k(g|w_{-k}^{-1}) = \inf \{ p(g|z_{-\infty}^{-1}) : z_{-\infty}^{-1} \in \mathcal{H}, z_{-k}^{-1} = w_{-k}^{-1} \}, \quad g \in G, \quad w_{-\infty}^{-1} \in \mathcal{H}, \quad (11)
\]
and the increments \( b_k : G \times \mathcal{H} \to [0,1] \) defined as
\[
b_k(g|w_{-k}^{-1}) = a_k(g|w_{-k}^{-1}) - a_{k-1}(g|w_{-(k-1)}^{-1}),
\]
for any \( k \in \mathbb{N} \), with \( a_{-1} \equiv 0 \).

In order to define the maximal coupling function we need to assume that, for any \( g \in G \), the function \( w_{-\infty}^{-1} \in \mathcal{H} \mapsto p(g|w_{-k}^{-1}) \) is continuous, i.e.
\[
\sum_{k=1}^{n} b_k(g|w_{-k}^{-1}) = a_n(g, w_{-n}^{-1}) \uparrow p(g|w_{-\infty}^{-1}), \quad \forall g \in G, \quad w_{-\infty}^{-1} \in \mathcal{H}. \quad (12)
\]
Also define, for any \( w_{-\infty}^{-1} \in \mathcal{H} \) and \( k \in \mathbb{N} \)
\[
a_k(w_{-k}^{-1}) := \sum_{g \in G} a_k(g|w_{-k}^{-1}). \quad (13)
\]
It is easily proved that (12) is equivalent to
\[
a_k(w_{-k}^{-1}) \uparrow 1, \quad \forall w_{-\infty}^{-1} \in \mathcal{H}. \quad (14)
\]
Next we partition the interval \([a_{k-1}(w_{-k+1}^{-1}), a_k(w_{-k}^{-1}))\) in subintervals \( B_k(g|w_{-k}^{-1}) \) of length \( b_k(g|w_{-k}^{-1}) \) (if this value is positive), varying \( g \in G \), for any \( k \in \mathbb{N} \): in the union \( \cup_{k=0}^{\infty} B_k(g|w_{-k}^{-1}) \) the function \( f(u|w_{-\infty}^{-1}) \) takes the value \( g \). Any function of this form will be called a maximal coupling function.

Also in the definition of the maximal coupling function it is possible to replace the set \( \mathcal{H} \) with a smaller set \( \mathcal{H}' \), provided the assumptions of Proposition 2 are satisfied. We will see a particular example in the sequel.

3 Backward coalescence times constructed from information depths

In this section we present a particular class of backward coalescence times specified through a two-stage procedure. This concept is inspired by the particular construction of the backward coalescence time presented in [CFF02].
An information depth \( K_0 = K_0(U_0^\infty) \) for the coupling function \( f \) is an a.s. finite stopping time w.r.t. the filtration \( \{ F_n^0, n \in \mathbb{N} \} \) with the property that \( K_0 = m \) implies

\[
\begin{align*}
f^{(m+1)}(U_0, U_{-1}, \ldots, U_{-m}|w_{-\infty}^{-(m+1)}) &= f(U_0)f^{(m)}(U_{-1}, \ldots, U_{-m}|w_{-\infty}^{-(m+1)}) \cdots f^{(1)}(U_{-m}|w_{-\infty}^{-(m+1)}), w_{-\infty}^{-(m+1)}) \\
&= f(U_0)f^{(m)}(U_{-1}, \ldots, U_{-m}|w_{-\infty}^{-(m+1)}) \cdots f^{(1)}(U_{-m}|w_{-\infty}^{-(m+1)}), g_{-\infty}^{-(m+1)}),
\end{align*}
\]

for any \( m \in \mathbb{N} \) and any \( w_{-\infty}^{-(m+1)} \in \mathcal{H} \), such that

\[
(f^{(m)}(U_{-1}, \ldots, U_{-m}|w_{-\infty}^{-(m+1)}), \ldots, f^{(1)}(U_{-m}|w_{-\infty}^{-(m+1)}), g_{-\infty}^{-(m+1)}) \in \mathcal{H}.
\]

It is checked that when this is fulfilled it remains true for any \( m > K_0 \). In fact the set of equalities (15) which have to be checked, for a fixed \( w_{-\infty}^{-(m+1)} \in \mathcal{H} \), is reduced as \( m \) grows. We recall that \( g_{-\infty}^{-(m+1)} \in \mathcal{H} \) is arbitrary, so if \( m \geq K_0 \), the dependence of the coupling function \( f^{(m+1)}(U_0, U_{-1}, \ldots, U_{-m}|w_{-\infty}^{-(m+1)}) \), on the history \( w_{-\infty}^{-(m+1)} \in \mathcal{H} \) prior to time \(-m_1\), is due only to the states computed in the subsequent interval \([-m, -1]\). For \( m = 0 \) property (15) means that \( f(U_0|w_{-\infty}^{-(m+1)}) \) is constant w.r.t. \( w_{-\infty}^{-(m+1)} \in \mathcal{H} \).

By comparing (15) with (7) it is seen that \( K_0 \) is not necessarily the negative of a backward coalescence time. In general, to eliminate completely the dependence on \( w_{-\infty}^{-(m+1)} \in \mathcal{H} \) in (15), a larger value of \( m \) has to be expected. In order to construct a backward coalescence time we define the sequence

\[
K = \{ K_j = K_0(U_j^\infty), j \in \mathbb{Z} \},
\]

of information depths at all times. Next introduce the random variable

\[
\tau_0^K(U_0^\infty) = \sup \{ s \leq 0 : K_j \leq j - s, s \leq j \leq 0 \}.
\]

Notice that, differently from \( K_0 \), \( \tau_0^K \) takes negative values: indeed, by definition \( \tau_0^K \leq -K_0 \). The random variable \( \tau_0^K \) is a candidate for a backward coalescence time; in fact the following result holds.

**Proposition 3.** If \( \tau_0^K \) is a.s. finite it satisfies properties H1 and H2.

*Proof.* Let us observe that, for any \( m \in \mathbb{N} \),

\[
\{-\tau_0^K \leq m\} = \bigcup_{i=0}^{m} F_i,
\]

where

\[
F_i = \{ K_{-i} = 0, K_{-i+1} \leq 1, \ldots, K_{-1} \leq i - 1, K_0 \leq i \}.
\]

Since \( F_i \in F_{-i}^0 \), H1 is proved.
Now assume that \( F_i \) is realized, for some \( i \in \mathbb{N} \). From \( K_{-i} = 0 \) it is obtained that \( F_i \) implies
\[
f(U_{-i}|w_{-\infty}^{-\langle i+1 \rangle}) = f(U_{-i}|g_{-\infty}^{-\langle i+1 \rangle}),
\]
for any \( w_{-\infty}^{-\langle i+1 \rangle} \in \mathcal{H} \), and thus
\[
f^{(2)}(U_{-i+1}, U_{-i}|g_{-\infty}^{-\langle i+1 \rangle}) = f(U_{-i+1}|f(U_{-i}|g_{-\infty}^{-\langle i+1 \rangle}), g_{-\infty}^{-\langle i+1 \rangle}) = f(U_{-i+1}|f(U_{-i}|w_{-\infty}^{-\langle i+1 \rangle}), g_{-\infty}^{-\langle i+1 \rangle})
\]
for any \( w_{-\infty}^{-\langle i+1 \rangle} \in \mathcal{H} \). Using \( K_{-i+1} \leq 1 \) and \((15)\) we obtain that
\[
f(U_{-i+1}|f(U_{-i}|w_{-\infty}^{-\langle i+1 \rangle}), g_{-\infty}^{-\langle i+1 \rangle}) = f(U_{-i+1}|f(U_{-i}|w_{-\infty}^{-\langle i+1 \rangle}), w_{-\infty}^{-\langle i+1 \rangle}) = f^{(2)}(U_{-i+1}, U_{-i}|w_{-\infty}^{-\langle i+1 \rangle}).
\]

By induction, using the same argument, it is obtained that
\[
f^{(i)}(U_0, \ldots, U_{-i+1}, U_{-i}|w_{-\infty}^{-\langle i+1 \rangle}) = f^{(i)}(U_0, \ldots, U_{-i+1}, U_{-i}|g_{-\infty}^{-\langle i+1 \rangle}),
\]
for any \( w_{-\infty}^{-\langle i+1 \rangle} \in \mathcal{H} \). Since \( K_{-i} \leq m \) means that \( F_i \) is realized for some \( i \in [0, m] \) and the property \((22)\) is preserved for values larger than \( i \), we get that H2 is fulfilled, too. \( \Box \)

Under the assumption of Proposition 3 we can define the process \( X \) through \((7)\) and \((10)\), with the shifted backward coalescence times defined by
\[
\tau_n = \tau_n^K = \sup\{ s \leq n : K_j \leq j - s, s \leq j \leq n \},
\]
for any \( n \in \mathbb{Z} \). Since \( m \in [\tau_n^K, n] \) implies \( \tau_m^K \geq \tau_n^K \), by starting the forward simulation from time \( \tau_n^K \), it is possible to recover all the values \( X_m \), with \( m = \tau_n^K, \ldots, n \), through the iteration of the coupling function with the (arbitrary) initial history \( g_{-\infty}^{-\langle \tau_n^K-1 \rangle} \).

Next we introduce a property which is stronger than the a.s. finiteness of \( \tau_0^K \), but easier to verify; this is again suggested by \( \text{[CFF02]} \). In order to introduce it, define the event
\[
R^K = \{ \tau_n^K \geq 0, \ \forall n \in \mathbb{N} \} = \{ \tau_0^K[0, \infty] = 0 \},
\]
where \( \tau_0^K[0, \infty] = \inf\{ \tau_i^K : i \geq 0 \} \). \( R^K \) belongs to the \( \sigma \)-algebra \( \mathcal{F}_0^{+\infty} = \sigma(\cup_{n \in \mathbb{N}} \mathcal{F}_0^n) \). When \( R^K \) is realized the iteration of the coupling function, started at time 0 from the arbitrarily chosen history \( g_{-\infty}^{-1} \in \mathcal{H} \), produces \( X_n \) for any \( n \in \mathbb{N} \).

It is convenient to rephrase the event \( R^K \) in terms of the information depths \( \{ K_n, n \in \mathbb{N} \} \). This is done in the following simple lemma, which is essentially borrowed from \( \text{[CFF02]} \).

**Lemma 1.** The following identity holds
\[
R^K = \{ K_n \leq n, n \in \mathbb{N} \}. \tag{24}
\]
Proof. First we prove that \( R^K \) is included in the r.h.s. of (24). Observe that, for any \( n \in \mathbb{N} \), it is \( K_j \leq j - \tau^K_n \) for \( j \in [\tau^K_n, n] \); in particular, for \( j = n \), \( K_n \leq n - \tau^K_n \). This does not exceed \( n \) provided \( \tau^K_n \geq 0 \), which proves the promised inclusion. For the converse inclusion the argument is the following. If \( K_n \leq n \) for any \( n \in \mathbb{N} \), then it is seen from (23) that 0 belongs to the set whose supremum is precisely \( \tau^K_n \). Thus \( \tau^K_n \geq 0 \), for any \( n \in \mathbb{N} \). \( \square \)

**Proposition 4.** If \( P(R^K) > 0 \) then \( \tau^K[0, \infty] \) is finite a.s.

Proof. Define the sequence of events \( \{ R^K_m, m \in \mathbb{Z} \} \) as

\[
R^K_m = \{ \tau^K_{m+n} \geq m, n \in \mathbb{N} \} = \{ K_{m+l} \leq l, l \in \mathbb{N} \}.
\]

Without loss of generality, working with the canonical realization of the process \( U \), we can identify \( R^K_m \), for any \( m \in \mathbb{Z} \), as an event of the sequence space \( [0,1)^\mathbb{Z} \). Then \( R^K_m = \sigma^m(R^K) \), where \( \sigma \) is the forward unit shift on the sequence space \( [0,1)^\mathbb{Z} \). By using the ergodic theorem it is deduced that there exists an a.s. finite \( Y_0 \leq 0 \) such that \( R^K_{R^0} \) is realized. Since \( 0 \geq \tau^K[0, \infty] \geq \tau^K[Y_0, \infty] \) = \( Y_0 \), the former is a.s. finite. \( \square \)

If \( P(R^K) > 0 \), along the same lines of the proof of Corollary 4.3 in [CFF02], it is possible to prove also that \( \sum_{k \in \mathbb{Z}} \delta_{R^k_k} \) is a stationary renewal process.

In principle, by starting the computation of the coupling function at time \( Y_0 \) (from the arbitrarily chosen history \( g_{Y_0-1} \)), we can construct the process \( X_n \) for all times \( n \geq Y_0 \). However notice that, \( -Y_0 \) being a stopping time w.r.t. the filtration \( \{ F_m^\pm; m \in \mathbb{N} \} \), it is not accessible by simulation.

We finally describe the construction in [CFF02] of an information depth for the maximal coupling function given in the previous section, with the adjustment needed to take into account its dependence on the admissible histories. Let us define

\[
a_k = \inf \left\{ a_k \left( w_k \right)^{-1} : w_k^{-1} \in \mathcal{H}\right\}, \quad k \in \mathbb{N}_+,
\]

where \( a_k \left( w_k \right)^{-1} \) has been defined in (13), and assume that the non decreasing sequence \( \{ a_k, k \in \mathbb{N} \} \) tends to 1, as \( k \to \infty \). This guarantees that \( a_k(g|w_k^{-1}), k \in \mathbb{N} \} \) converges to \( p(g|w_k^{-1}) \), for any \( w_k^{-1} \in \mathcal{H} \) and any \( g \in G \), as \( k \to \infty \): by consequence the maximal coupling function \( f(\cdot|w_k^{-1}) \) is well defined for any \( w_k^{-1} \in \mathcal{H} \). Next define \( K : [0,1) \to \mathbb{N} \) as

\[
K(u) = \sum_{k=0}^{\infty} k 1_{[a_{k-1}, a_k)}(u) = \inf\{k \in \mathbb{N} : a_k > u\} \quad (25)
\]

where \( a_{-1} = 0 \). Since \( a_k \leq a_k(w_k^{-1}) \), for any \( w_k^{-1} \in \mathcal{H} \), if \( K(u) = k \), the coupling function \( f(u|w_k^{-1}) \) is computable by knowing only \( w_k^{-1} \). Thus if \( a_k \uparrow 1 \), as \( k \to \infty \), condition (15) holds for \( K_0 = K(U_0) \), hence \( K(U_0) \) is an information depth.

Sufficient conditions on the sequence \( \{ a_k, k \in \mathbb{N} \} \) which guarantee that the stopping times \( \tau^K_0 \) and \( \tau^K[0, \infty] \) (with \( K_j = K(U_j) \)), are finite a.s., can be found in [CFF02]. We summarize the main results in the following proposition.
Proposition 5. Let \( \{U_i, i \in \mathbb{N}\} \) be a sequence of i.i.d. random variables, uniformly distributed in the interval \([0,1)\), and let \( \{a_k \in [0,1], k \in \mathbb{N}\} \) be a sequence increasing to 1. Define \( K \) as in (25).

(a) If \( \sum_{k=0}^{\infty} \prod_{j=0}^{k} a_j = \infty \) (which implies \( a_k \uparrow 1, \) as \( k \to \infty \)), then

\[
\tau_0^K = \sup \{ s \leq 0 : K(U_j) \leq j - s, s \leq j \leq 0 \} > -\infty, \text{ a.s.}
\]

(b) If \( \prod_{j=0}^{\infty} a_j > 0 \), then \( P(R_{K}) > 0 \), thus

\[
\tau^K[0,\infty] = \sup \{ s \leq 0 : K(U_j) \leq j - s, s \leq j \} > -\infty, \text{ a.s.}
\]

4 An information depth depending on the whole past

In order to motivate the present section we start with a class of examples for which the sufficient conditions of Proposition 5 appear to be rather restrictive.

Example 1. Assume \( G = \{-1,1\} \), and consider a kernel \( p : G \times G^{-N^*} \to [0,1] \) of the following form

\[
p(j|w_{-\infty}^{-1}) = \begin{cases} 
p_k(w_{-1}, j), & \text{if } w_{-1} = \cdots = w_{-k}, \ w_{-k-1} = -w_{-1}, \\
p_{\infty}(w_{-1}, j), & \text{if } w_{-n} = w_{-1}, \ n \geq 1,
\end{cases}
\]

where

\[
P_k = \{ p_k(i, j) : i, j \in \{-1,+1\}\}
\]

is a stochastic matrix for any \( k \in \mathbb{N}^* \cup \{\infty\} \). The value \( p_k(i, j) \) gives the conditional probability that the next state is equal to \( j \) given that the previous \( k \) states are all equal to the current state \( i \), and the \((k+1)\)-th equals \(-i\) (when \( k < \infty \)). Since, after the first change of sign, further information about the past is not relevant, it appears that \( p_k(i, i), i \in \{-1,+1\} \) represents the survival rates of an alternating renewal process (see [BB03] pp. 32-35).

We assume that there exists \( \epsilon > 0 \) such that, for \( i = -1,1 \), it holds

\[
\epsilon \leq p_h(i, i) \leq 1 - \epsilon, \ i \in \mathbb{N}^* \cup \{\infty\}.
\]

In this case obviously \( H = G^{-N^*} \); however it is possible to consider, in the definition of the maximal coupling function, the smaller set of histories

\[
\mathcal{H}' = \{ w_{-\infty}^{-1} : \sum_{i=1}^{\infty} \delta_{w_{-i},1} = +\infty, \sum_{i=1}^{\infty} \delta_{w_{-i},-1} = +\infty \}.
\]
The assumptions of Proposition 2 are rather trivially checked. 1. holds since the occurrence of \( w_{-\infty}^1 \in \mathcal{H} \) depends on the tail of the sequence \( w_{-\infty} \). 2. is due to the fact that \( f(u) \) is constant on sufficiently small balls in \( \mathcal{H} \). Finally 3. and 4. are ensured by (26).

Moreover
\[
a_0(-1) = \min \{ \inf_{h \in \mathbb{N}_+} p_h(-1, -1), 1 - \sup_{h \in \mathbb{N}_+} p_h(1, 1) \} \geq \epsilon, \\
a_0(1) = a_0 - a_0(-1) = \min \{ 1 - \sup_{h \in \mathbb{N}_+} p_h(-1, -1), \inf_{h \in \mathbb{N}_+} p_h(1, 1) \} \geq \epsilon.
\]

As a consequence
\[
\tau_0(U_{-\infty}^0) = \sup \{ n \in -\mathbb{N} : 1_{[a_0(-1), a_0)}(U_n)1_{[a_0(-1), a_0)}(U_{n-1}) + 1_{[a_0(-1), a_0)}(U_{n-1})1_{[0, a_0(-1))}(U_n) = 1 \}
\]

is a.s. finite and it is readily checked to be a backward coalescence time. Applying Proposition 2 the existence and uniqueness of a stationary alternating renewal process is proved. Condition (26) is certainly more restrictive than needed, since it implies that the distributions of the holding times have exponential tail bounds, whereas it is enough that they have finite mean, see [BB03].

Next we check the sufficient conditions in Proposition 5. Since
\[
a_k = 1 - \max(\sup_{h \geq k} p_h(-1, -1) - \inf_{h \geq k} p_h(-1, -1), \sup_{h \geq k} p_h(1, 1) - \inf_{h \geq k} p_h(1, 1)), k \in \mathbb{N}_+,
\]

it follows that \( a_k \to 1 \), as \( k \to \infty \), if and only if \( p_k(i, i) \) converge as \( k \to \infty \), for \( i = -1, 1 \). If this condition fails Proposition 5 cannot be applied. Even if this condition holds, the speed of convergence of \( a_k \) to 1 can be so slow that condition (a) in Proposition 5 is still violated. This happens, for example, for
\[
p_h(i, i) = \frac{1}{2} \left( 1 - \frac{1}{\sqrt{h + 1}} \right), \quad i \in \{-1, 1\}, h \in \mathbb{N}_+ \cup \{\infty\}.
\]

The previous example suggests to investigate alternative ways to define an information depth for the maximal coupling function, order for the construction of a backward coalescence time under weaker assumptions than those given in Proposition 5.

Let us introduce the process \( \{A_h, h \in \mathbb{N}\} \), with \( A_0 = a_0 \) and
\[
A_h(U_{-1}, \ldots, U_{-h}) := \inf \{ a_h(w_{-h}) : w_{-\infty} \in J_h(U_{-1}, \ldots, U_{-h}) \}, \quad h \in \mathbb{N}_+,
\]
where
\[
J_h(U_{-1}, \ldots, U_{-h}) = \{ w_{-\infty} \in \mathcal{H} : w_{-k} = g, \text{ if } U_{-k} \in B_0(g), \text{ } g \in G, \text{ } k \leq h \}
\]
is a set of histories smaller than \( \mathcal{H} \), since the states which can be identified, given \( U_{-1}, \ldots, U_{-h} \), are kept fixed. In fact, whenever \( U_{-k} \in B_0(g) \), the function \( f(U_{-k}; w_{-\infty}^{-(k+1)}) \) is equal to \( g \),
irrespectively of the previous history $w_{-\infty}^{-(k+1)} \in \mathcal{H}$. Since $A_h(U_{-1}, \ldots, U_{-h})$ is an infimum computed on a smaller set, it holds

$$A_h(U_{-1}, \ldots, U_{-h}) \geq a_h, h \in \mathbb{N},$$

(30)

for any realization of the i.i.d. sample $U_{-h}^{-1}$.

Now let us define

$$K' (U_{-\infty}^0) = \inf\{ j \in \mathbb{N} : U_0 < A_j(U_{-1}, \ldots, U_{-j}) \}. \quad (31)$$

**Proposition 6.** If $\lim_h A_h(U_{-1}, \ldots, U_{-h}) = 1$ a.s., $K'$ is an information depth for the maximal coupling function.

**Proof.** Only property (15) needs some explanations. Suppose that $K'(U_{-\infty}^0) = m$. Then

$$U_0 < A_m(U_{-1}, \ldots, U_{-m}) \leq a_m(w_{-m}^{-1}), \forall w_{-m}^{-1} \in J_m(U_{-1}, \ldots, U_{-m}),$$

which from the definition of the maximal coupling function implies

$$f(U_0|w_{-\infty}^{-1}) = f(U_0|z_{-\infty}^{-1}), \text{ for } w_{-\infty}^{-1}, z_{-\infty}^{-1} \in J_m(U_{-1}, \ldots, U_{-m}), \text{ such that } w_{-m}^{-1} = z_{-m}^{-1}. \quad (32)$$

Finally consider any $w_{-\infty}^{-(m+1)} \in \mathcal{H}$; then, by choosing

$$w_{-m}^{-1} = (f^{(m)}(U_{-1}, \ldots, U_{-m}|w_{-\infty}^{-(m+1)}), \ldots, f^{(1)}(U_{-m}|w_{-\infty}^{-(m+1)})),$$

the history $w_{-\infty}^{-1} \in J_m(U_{-1}, \ldots, U_{-m})$. Now set $z_{-\infty}^{-1} = (w_{-m}^{-1}, g_{-\infty}^{-(m+1)})$. If $z_{-\infty}^{-1} \in \mathcal{H}$ then it belongs necessarily to $J_m(U_{-1}, \ldots, U_{-m})$, in which case formula (32) is turned into (15).

\[\square\]

Next define $\{K'_j = K'(U_{-\infty}^j), j \in \mathbb{Z}\}$, and $\tau_0^{K'}$ and $R_0^{K'}$ as in (18) and (21), respectively: then the following result holds.

**Theorem 1.** If $\prod_{h=0}^{\infty} A_h(U_{h-1}, \ldots, U_0)^{-1} \in \mathcal{L}^1$ then $P(R_0^{K'}) > 0$.

**Proof.** The sequence

$$Y_n := \frac{\prod_{h=0}^n 1\{U_h < A_h(U_{h-1}, \ldots, U_0)\}}{\prod_{h=0}^n A_h(U_{h-1}, \ldots, U_0)}, \quad n \in \mathbb{N},$$

(33)

is uniformly integrable, because it is bounded by the integrable random variable $\prod_{h=0}^{\infty} A_h(U_{h-1}, \ldots, U_0)^{-1}$. Moreover it is a martingale with respect to the filtration $\{\mathcal{F}_0^n = \sigma(U_n, \ldots, U_0) : n \in \mathbb{N}\}$. In fact, since $U_n$ is independent of $\mathcal{F}_0^{n-1}$, it holds

$$E\left(1\{U_n < A_n(U_{n-1}, \ldots, U_0)\} | \mathcal{F}_0^{n-1}\right) = A_n(U_{n-1}, \ldots, U_0),$$
hence

\[
E(Y_n | \mathcal{F}_0^{n-1}) = \frac{\prod_{h=0}^{n-1} \mathbf{1}_{\{U_h < A_h(U_{h-1}, \ldots, U_0)\}}}{\prod_{h=0}^{n} A_h(U_{h-1}, \ldots, U_0)} E \left( \mathbf{1}_{\{U_n < A_n(U_{n-1}, \ldots, U_0)\}} | \mathcal{F}_0^{n-1} \right) = Y_{n-1}.
\]

Furthermore, since \( Y_0 = a_0^{-1} \mathbf{1}_{\{U_0 < a_0\}} \), it follows that \( E(Y_n) = E(Y_0) = 1 \).

Since \( \{Y_n\} \) is uniformly integrable, from a well known result (see [Wil91], page 134), the limit \( Y_\infty := \lim_{n \to \infty} Y_n \) is finite a.s. with \( E(Y_\infty) = 1 \). As a consequence \( Y_\infty > 0 \) with positive probability. But clearly

\[
\{Y_\infty > 0\} = \bigcap_{h=0}^{\infty} \{U_h < A_h(U_{h-1}, \ldots, U_0)\} = \{K'_h \leq h, h \in \mathbb{N}\} = R_0^{K'},
\]

which ends the proof.

By applying Proposition 4, the previous theorem implies that a perfect simulation algorithm can be constructed from the information depths \( \{K'_j, j \in \mathbb{Z}\} \).

**Example 1** (continued). We prove that Theorem 1 can be applied to Example 1, under the assumption

\[
a_0(-1) > 0, \quad a_0(1) > 0, \quad a_\infty := \sup_{k \geq 0} a_k > 1 - 2a_0(-1)a_0(1).
\]

Notice that the latter condition is automatically verified if the former ones hold and \( a_\infty = 1 \). Define

\[
N(U_0^\infty) = \inf \{n \in \mathbb{N} : 1_{[a_0(-1), a_0]}(U_n)1_{[0, a_0(-1)]}(U_{n-1}) + 1_{[a_0(-1), a_0]}(U_{n-1})1_{[0, a_0(-1)]}(U_n) = 1\}.
\]

Then if \( h > N(U_0^\infty) \), and \( w_{-\infty}^{h-1} \in J_h(U_{h-1}, \ldots, U_0) \) then

\[
a_h(g|w_{-\infty}^{h-1}) = p(g|w_{-\infty}^{h-1}).
\]

In this case

\[
A_h(U_{h-1}, \ldots, U_0) = \inf \left\{ \sum_{g \in G} a_h(g|w_{0}^{h-1}) : w_{-\infty}^{h-1} \in J_h(U_{h-1}, \ldots, U_0) \right\}
\]

\[
= \inf \left\{ \sum_{g \in G} p(g|w_{-\infty}^{h-1}) : w_{-\infty}^{h-1} \in J_h(U_{h-1}, \ldots, U_0) \right\} = 1.
\]
Therefore, for any $\delta \in (0, a_{\infty} - 1 + 2a_0(-1)a_0(1))$, choosing $n_0 = n_0(\delta)$ such that $a_{n_0} \geq 1 - 2a_0(-1)a_0(1) + \delta$, it is obtained

$$
\left( \prod_{h=0}^{\infty} A_h(U_{h-1}, \ldots, U_0) \right)^{-1} = \left( \prod_{h=0}^{N} A_h(U_{h-1}, \ldots, U_0) \right)^{-1} \leq (1 - 2a_0(-1)a_0(1) + \delta)^{-N} \prod_{h=0}^{n_0} a_h^{-1}.
$$

(37)

The expression (36) suggests a majorization of $N$ with twice a geometric random variable having the success probability $p = 2a_0(1)a_0(-1)$. Since the radius of convergence for the p.g.f. of this kind of random variable is $1/(1 - p)$ the assumption (35) and the bound (37) imply that $\prod_{h=0}^{\infty} A_h(U_{h-1}, \ldots, U_0)^{-1}$ is integrable. Therefore Theorem 1 can be applied, showing that the stationary alternating renewal process can be perfectly simulated.

Inspired by the previous example, in the following corollary we present a sufficient condition, possibly easier to verify, which guarantees that the assumption in Theorem 1 holds.

**Corollary 1.** Let $C_h$ be a Borel subset of $[0, 1)^h$ and suppose:

1) the sequence

$$
\alpha_h = \inf \{ A_h(u_{h-1}, \ldots, u_0) : (u_{h-1}, \ldots, u_0) \in C_h \}
$$

is such that $\prod_{h=0}^{\infty} \alpha_h > 0$ (in particular $\alpha_0 = a_0 > 0$);

2) the random variable

$$
N := N(U_{0}^{+\infty}) = \inf \{ m : (U_{n-1}, \ldots, U_0) \in C_n, \forall n \geq m \}
$$

has a probability generating function $E(s^N) < \infty$ for some $s > 1/a_{\infty}$.

Then

$$
E \left( \prod_{h=0}^{\infty} A_h(U_{h-1}, \ldots, U_0)^{-1} \right) < \infty.
$$

(40)

Proof. By definition of $N$, $A_h \geq \alpha_h$ for $h > N$. Moreover, since $A_h \geq a_h$ for each integer $h$, we have

$$
\frac{1}{\prod_{h=0}^{\infty} A_h} \leq \frac{1}{\prod_{h=0}^{N} a_h} \frac{1}{\prod_{h=N+1}^{\infty} \alpha_h}.
$$

(41)

By taking expected values at both sides, with a straightforward bound for the second factor at the r.h.s., it is obtained

$$
E \left( \frac{1}{\prod_{h=0}^{\infty} A_h} \right) \leq E \left( \frac{1}{\prod_{h=0}^{N} a_h} \right) \frac{1}{\prod_{h=0}^{\infty} \alpha_h}.
$$

(42)
The second factor at the r.h.s. is finite by assumption 1). By assumption there exists an integer \( k \) such that \( \frac{1}{a_k} < s \), \( s \) being as in 2). By consequence we have the following bound for the first factor

\[
E \left( \frac{1}{\prod_{h=0}^{N} a_h} \right) \leq \frac{1}{\prod_{h=0}^{k-1} a_h} E(s^N) < +\infty,
\]

(43)

from which the corollary follows.

Finally we provide another class of models that satisfy the conditions of Corollary 1 but not those of Proposition 5.

**Example 2.** Consider positive summable sequences \( \beta(i) \geq \gamma(i), \ i \in \mathbb{N}_+ \), and assume that \( p_1 \in (0, 1), \sigma > 0 \) and \( c > 0 \) are such that

\[
p_1(1 - c \sum_{i=1}^{\infty} \beta(i)) > \sigma;
\]

(44)

moreover assume that \( \sum_{i=1}^{\infty} i \gamma(i) < \infty \).

Now define the kernel \( p \) on \( G = \{0, 1\} \) by

\[
p(1|\mathbf{w}_{-\infty}^{-1}) = p_1 \left\{ 1 - c \sum_{i=1}^{\infty} (\beta(i) 1_{\{w_{-i,0}, T(\mathbf{w}_{-\infty}^{-1}) > i\}} + \gamma(i) 1_{\{w_{-i,0}, T(\mathbf{w}_{-\infty}^{-1}) \leq i\}}) \right\},
\]

where

\[
T(\mathbf{w}_{-\infty}^{-1}) = \inf \left\{ k : \frac{\sum_{i=1}^{k} w_{-i}}{k} \geq \sigma \right\}.
\]

(45)

First of all we prove that the kernel \( p \) is monotone, which means that \( p(1|\mathbf{w}_{-\infty}^{-1}) \) is increasing in \( \mathbf{w}_{-\infty}^{-1} \) w.r.t. the pointwise order. For this notice that \( w_{-i} \geq \eta_{-i} \), for \( i \in \mathbb{N}_+ \) implies \( T(\mathbf{w}_{-\infty}^{-1}) \leq T(\eta_{-\infty}^{-1}) \), hence

\[
p(1|\mathbf{w}_{-\infty}^{-1}) \geq p_1 \left\{ 1 - c \sum_{i=1}^{\infty} (\beta(i) 1_{\{\eta_{-i,0}, T(\mathbf{w}_{-\infty}^{-1}) > i\}} + \gamma(i) 1_{\{\eta_{-i,0}, T(\mathbf{w}_{-\infty}^{-1}) \leq i\}}) \right\}
\]

\[
\geq p_1 \left\{ 1 - c \sum_{i=1}^{\infty} (\beta(i) 1_{\{\eta_{-i,0}, T(\eta_{-\infty}^{-1}) > i\}} + \gamma(i) 1_{\{\eta_{-i,0}, T(\eta_{-\infty}^{-1}) \leq i\}}) \right\} = p(1|\eta_{-\infty}^{-1}),
\]

where the second inequality is due to the fact that \( \beta(i) \geq \gamma(i) \), for \( i \in \mathbb{N}_+ \).

Since \( a_0(1) = p_1(1 - c \sum_{i=1}^{\infty} \beta(i)) > 0 \) and \( a_0(0) = 1 - p_1 > 0 \) it follows that \( \mathcal{H} = G^{-\mathbb{N}_+} \).

As a consequence

\[
a_k(0, \mathbf{w}_{-k}^{-1}) = \inf \left\{ p(0|\mathbf{w}_{-k}^{-1}, \mathbf{z}_{-\infty}^{-k-1}) : \mathbf{z}_{-\infty}^{-k-1} \in \{0, 1\}^{-\mathbb{N}_+} \right\} = p(0|\mathbf{w}_{-k}^{-1}, 1_{-\infty}^{-k-1})
\]

(47)
and
\[ a_k(1, w^{-1}_k) = \inf \{ p(1 \mid w^{-1}_k, z^{-k-1}_\infty) : z^{-k-1}_\infty \in \{0, 1\}^{-N_+} \} = p(1 \mid w^{-1}_k, 0^{-k-1}) \] (48)

therefore
\[ a_k(w^{-1}_k) = p(0 \mid w^{-1}_k, 1^{-k-1}_\infty) + p(1 \mid w^{-1}_k, 0^{-k-1}) = 1 - p(1 \mid w^{-1}_k, 1^{-k-1}_\infty) + p(1 \mid w^{-1}_k, 0^{-k-1}) \] (49)

which by a direct computation is seen to assume only the value \( s_1 - p_1 c \sum_{i=k+1}^{\infty} \gamma(i) \), when \( T(w^{-1}_\infty) \leq k \), and \( 1 - p_1 c \sum_{i=k+1}^{\infty} \beta(i) \), otherwise. Notice that the condition \( T(w^{-1}_\infty) \leq k \) can be verified by looking only at \( w^{-1}_k \).

Now we prove that this class of kernels can be perfectly simulated. In fact we can prove that the conditions given in Corollary 1 are satisfied for the sequence of events
\[ C_n = \left\{ (u_{n-1}, \ldots, u_0) \in [0, 1]^n : \frac{1}{n} \sum_{k=0}^{n-1} 1_{\{a_0(0) \leq u_k < a_0(0) + a_0(1)\}} \geq \sigma. \right\} \] (50)

Since \( a_0(1) > \sigma \), from Chernoff’s bound
\[ P((U_{n-1}, \ldots, U_0) \notin C_n) \leq e^{-Kn}, \]
for some \( K > 0 \), therefore
\[ P(N \leq n_0) = P(\cap_{n=n_0}^{\infty} \{(U_{n-1}, \ldots, U_0) \in C_n\}) \geq 1 - \sum_{n=n_0}^{\infty} e^{-Kn} = 1 - \frac{e^{-K n_0}}{1 - e^{-K}}, \]
from which the existence of the probability generating function of \( N \), for some \( s > 1 \), is deduced. From (51) and the summability of \( \beta(i) \), it is obtained that \( a_\infty = 1 \), which ensures that condition 2) of Corollary 1 is satisfied. Finally observe that whenever \((U_{n-1}, \ldots, U_0) \in C_n\)
\[ A_n(U_{n-1}, \ldots, U_0) \geq \inf \{ a_n(0 \mid w^{-1}_n) + a_n(1 \mid w^{-1}_n) : T(w^{-1}_\infty) \leq n \} = 1 - p_1 c \sum_{i=n+1}^{\infty} \gamma(i), \]
therefore, by definition (58), we get \( \alpha_n \geq 1 - p_1 c \sum_{i=n+1}^{\infty} \gamma(i) \). Now, being
\[ \sum_{n=0}^{\infty} \sum_{i=n+1}^{\infty} \gamma(i) = \sum_{i=1}^{\infty} i \gamma(i) < \infty \]
we get by [Wil91], page 40, that \( \prod_{n=1}^{\infty} \alpha_n > 0 \), so that condition 1) is also satisfied.
On the other hand, for some choices of \(\{\beta(i)\}\) and \(\{\gamma(i)\}\), condition a) in Proposition 5 fails. For example consider \(\beta(i) = i^{-\alpha}\), with \(\alpha \in (1, 2)\) and \(\gamma(i) = 2^{-i}\), which ensure that \(\beta(i) \geq \gamma(i)\), for \(i \in \mathbb{N}_+\), and \(\sum_{i=1}^{\infty} i \gamma(i) < \infty\). Since

\[
    a_k = \inf \{a_k(w_{-k}^{-1}) : w_{-k}^{-1} \in \{0, 1\}^k\} = 1 - p_1 c \sum_{i=k+1}^{\infty} \beta(i),
\]

we can show that \(\sum_{k=1}^{\infty} \prod_{i=1}^{k} a_i < \infty\). In fact

\[
    \sum_{k=1}^{\infty} \prod_{i=1}^{k} a_i = \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(1 - p_1 c \sum_{i=k+1}^{\infty} \frac{1}{i^{\alpha}}\right) \leq \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(1 - \frac{L_1}{i^{\alpha-1}}\right),
\]

where \(L_1 > 0\) is a sufficiently small constant. The rightmost expression is smaller than

\[
    \sum_{k=1}^{\infty} \exp\left(\sum_{i=1}^{k} -\frac{L_1}{i^{\alpha-1}}\right) \leq C \sum_{k=1}^{\infty} \exp(-L_2 k^{2-\alpha}) < \infty
\]

where \(C\) and \(L_2\) are suitable positive constants, which implies the promised inequality.

We conclude the section by observing that the idea of defining the information depth by computing the infimum of \(a_k(w_{-k}^{-1})\) over the set \(J_h(U_{-1}, \ldots, U_{-h})\) of histories compatible with the observed \(U_{-1}, \ldots, U_{-h}\) can be pushed further. For example, by looking at adjacent pairs \((U_{-i}, U_{-i+1}), i = 2, \ldots, h\), it is possible to locate other states, restricting the set of histories compatible with the observed \(U_{-1}, \ldots, U_{-h}\) to the smaller subset

\[
    J'_h(U_{-1}, \ldots, U_{-h}) = J_h(U_{-1}, \ldots, U_{-h}) \cap F_h(U_{-1}, \ldots, U_{-h}),
\]

where \(F_h(U_{-1}, \ldots, U_{-h})\) is equal to

\[
    \{w_{-\infty} \in \mathcal{H} : w_{-k} = g_1, w_{-k+1} = g_2, \text{ if } U_{-k} \in B_0(g_1), U_{-k+1} \in B_1(g_2|g_1), g_1, g_2 \in G, 2 \leq k \leq h\}.
\]

The changes to Proposition 6 and Theorem 1 are minor, but for the sake of brevity, we do not pursue this extension further.

5 An algorithm which works without minorization condition

In this section we explore the possibility of defining a backward coalescence time \(\tau_0\) when \(a_0 = 0\). In this case any information depth takes necessarily positive values, hence it
Proposition 7. If the random variable \( \tau_m \) from time which proves H1. Moreover, if this event is realized the process is in the markovian regime at time \( n \) whenever \( U_n < a_1 \). This means that the information needed to compute the state of the process at time \( n \) concerns only the state at time \( n - 1 \). For any \( u \in [0, 1) \) and \( w \in G \) we define

\[
\tilde{f}(u|w) := f(a_1u|w^{-1}_0),
\]

(52)

for any choice of \( w^{-1}_\infty \in \mathcal{H} \) having \( w^{-1}_1 = w \). Thus \( f : [0, 1) \times G \to G \), applied to a uniform random variable in \([0, 1]\), induces the Markov kernel

\[
M(g|w) = |\{u \in [0, 1) : \tilde{f}(u|w) = g\}|, \quad g, w \in G,
\]

(53)

where \(| \cdot |\) denote the Lebesgue measure.

By induction, for \( n \geq 2 \), we define the composition \( \tilde{f}^{(n)} : [0, 1)^n \times G \to G \) as

\[
\tilde{f}^{(n)}(u_n, \ldots, u_1|w) = \tilde{f}(u_n|\tilde{f}^{(n-1)}(u_{n-1}, \ldots, u_1|w)), \quad u_i \in [0, 1), \ i = 1, \ldots, n, \ w \in G
\]

(54)

where \( \tilde{f}^{(1)} = \tilde{f} \).

Concerning the markovian regime, for any \( n \in \mathbb{N}_+ \), we define the coalescence in the interval \([-n + 1, 0]\) as

\[
E_n = \{(u_0, u_{-1}, \ldots, u_{-n+1}) \in [0, 1)^n : \tilde{f}^{(n)}(u_0, \ldots, u_{-n+1}|w) = \tilde{f}^{(n)}(u_0, \ldots, u_{-n+1}|g_0), \forall w \in G\},
\]

(55)

where \( g_0 \in G \) is an arbitrary state.

We notice that if the kernel \( p \) is markovian, then \( a_1 = 1 \) and conversely. In this case \( \tilde{f} = f \) and any backward coalescence time has the property that \( \tau_0 = -m \) implies that \( (U_0, \ldots, U_{-m+1}) \in E_m \), as in the original CFTP algorithm [PW96].

Next assume that \( a_k \uparrow 1 \) as \( k \to \infty \) and recall that in this case \( K \), as defined in (25), takes finite values. Define the random variable \( \tau_0 \) as

\[
\sup\{m < 0 : \exists l \in [m, 0], \text{ s.t. } a_1^{-1}(U_l, \ldots, U_m) \in E_{l-m+1}, \text{ & } K(U_j) \leq j-l, j \in [l+1, 0] \text{ if } l < 0\}.
\]

(56)

**Proposition 7.** If the random variable \( \tau_0 \) is finite almost surely, it is a backward coalescence time.

**Proof.** By definition, for any \( m \in -\mathbb{N} \), the event \( \tau_0 = m \) belongs to the \( \sigma \)-algebra \( \mathcal{F}_m^0 \), which proves H1. Moreover, if this event is realized the process is in the markovian regime from time \( m \) to some larger time \( l \) in which coalescence has taken place. This means that

\[
\tilde{f}^{(l-m+1)}\left(\frac{U_l}{a_1}, \ldots, \frac{U_m}{a_1}|w\right) = \tilde{f}^{(l-m+1)}\left(\frac{U_l}{a_1}, \ldots, \frac{U_m}{a_1}|g_0\right)
\]
for any \( w \in G \). By the relation (52) this means that, for any \( w^{-1}_\infty \in H \), it holds

\[
f^{(l-m+1)}(U_{l}, \ldots, U_{m}|w^{-(m+1)}_\infty) = f^{(l-m+1)}(U_{l}, \ldots, U_{m}|g^{-(m+1)}_\infty)
\]  

(57)

Thus, if \( l = 0 \), H2 holds. If \( l < 0 \) one needs to repeat the proof of Proposition \( \ref{prop:relation} \) replacing (21) with (57). In short, to compute all the states of the process in the interval \([l + 1, 0]\), there is no requirement about the states of the process prior to time \( l \). \( \Box \)

After this result, we turn our interest to give sufficient conditions for the a.s. finiteness of \( \tau_0 \).

**Theorem 2.** Under the assumptions

(i) \( \sum_{n=1}^{\infty} \prod_{m=1}^{a} a_m = \infty \);

(ii) there exists \( s \in N_+ \) such that \( P((U_{s-1}, \ldots, U_0) \in E_s) > 0 \);

\( \tau_0 \) is finite a.s.

**Proof.** We start by defining the sequences \( \{W_n, n = 1, 2 \ldots\} \), \( \{Y_n, n = 1, 2, \ldots\} \) which will be proved to be finite a.s. First define

\[
W_1 = \sup\{m \leq 0 : K(U_j) - 1 \leq j - m, j \in [m, 0]\}.
\]

By Proposition \( \ref{prop:relation} \) part (a), condition (i) guarantees that \( W_1 \) is a.s. finite: notice indeed that replacing \( K(U_j) \) with \( K(U_j) - 1 \) has the effect of shifting the sequence \( \{a_j, j \in N\} \) to the left. Next define

\[
Y_i = \inf\{m < W_i : U_n < a_1, n \in [m+1, W_i]\}, \quad W_{i+1} = \sup\{m \leq Y_i : K(U_j) - 1 \leq j - m, j \in [m, Y_i]\},
\]

which are a.s. finite, for \( i \in N_+ \). It is immediately seen that \( \{W_i - Y_i - 1\}_{i \in N_+} \) is a sequence of i.i.d. geometric random variables, with success probability \( 1 - a_1 \). Likewise \( \{W_{i+1} - Y_i\}_{i \in N_+} \) is a sequence of i.i.d. random variables distributed as \( W_1 \), conditional to be non zero. Moreover the two sequences are mutually independent and independent of \( W_1 \). In particular the sequence \( \{-W_i, i \in N_+\} \) form a delayed renewal process and the sequence \( \{U_{-n}, n \in N\} \) is regenerative w.r.t. it.

Finally define the random index

\[
Q = \inf\{i \in N_+ : (U_{W_{i-1}}, \ldots, U_{Y_i+1}) \in E_{W_i-Y_i-1}\}
\]

(58)

and let \( \tau^* = Y_Q \).

For each \( j \in [W_i, 0] \) the condition \( j - K(U_j) \geq W_i - 1 \) is satisfied, for any \( i \in N_+ \). By consequence \( \tau^* \) differs from \( \tau_0 \) only because the supremum is taken on the set \( \{Y_i : i \in N_+\} \).
rather than on the whole set of negative integers. In fact notice that, for \( n = \tau^* = Y_Q \), one can always choose in (56) \( l = W_Q - 1 \). Therefore \( \tau^* \leq \tau_0 \), so it is enough to prove that \( \tau^* > -\infty \) a.s. But this is true because, by assumption (ii), the condition at the r.h.s. of (58) is fulfilled with positive probability in any regenerating cycle: an application of the law of large numbers concludes the proof.

\[ \square \]

In the previous theorem we have not assumed that \( a_0 = 0 \). However, in this case the result does not add anything to the statement (a) in Proposition 5. Indeed, if \( U_0 \) is uniformly distributed in \([0, 1)\), \( P(U_0 \in E_1) > 0 \) if \( a_0 > 0 \): assumption (ii) is always satisfied. Therefore, in the following we will always take \( a_0 = 0 \).

Assumption (ii) in the previous theorem states that the markovian coupling function \( \tilde{f} \) defined in (52) is successful for the perfect simulation of the Markov chain with kernel \( M \) given in (53), in the sense that backward coalescence occurs with probability 1. Since this implies the convergence in law of the chain as time increases, it is necessary that \( M \) has a single positive recurrent irreducible class which is aperiodic.

When \( G \) is finite, which is assumed from now on, this condition can be directly referred to the oriented graph induced by \( M \). Notice that if \((w, g)\) is an arc of this graph then necessarily \( a_1(g|w) > 0 \). However the converse is not true. In fact if \( B_1(g|w) \) is non empty and it is disjoint from \([0, a_1)\) then \((w, g)\) is not an arc. In this case, if \( w_{-1} = w \), the maximal coupling function can be replaced by a new coupling function \( \tilde{f}(u|w_{-\infty}) \) which is different only for \( u < a_1(w) \). Each interval \( B_1(h|w) \) is replaced by the union of two disjoint intervals \( B_1^1(h|w) \) and \( B_1^2(h|w) \), where \( \tilde{f}(.|w_{-\infty}) \) takes the value \( h \).

We require that

\[
|B_1^1(h|w)| + |B_1^2(h|w)| = |B_1(h|w)| = a_1(h|w) = |\{ u < a_1(w) : f(u|w_{-\infty}) = h \}| \quad (59)
\]

and \( B_1^1(h|w) \) intersects the interval \([0, a_1)\). Therefore the Markov kernel \( \tilde{M} \) induced by \( \tilde{f} \) satisfies

\[
a_1(g|w) > 0 \iff \tilde{M}(g|w) > 0. \quad (60)
\]

However, backward coalescence w.p. 1 cannot be ensured only by properties of the Markov kernel, without reference to the coupling function. A simple counterexample is presented in [Häg02]. But when the state space is finite, there is a universal modification of a Markov coupling function, which ensures backward coalescence w.p. 1 under the only assumption that the induced kernel has a unique irreducible class which is aperiodic (see Proposition 8.1 p. 122 in [AG07]). The modification consists in letting the different trajectories move independently before merging.

This is more clearly explained by allowing coupling functions to depend on \( n+1 \) variables \((u^0, \ldots, u^n) \in [0, 1)^{n+1}\), rather than a single variable \( u \in [0, 1)\); in the definition just replace
intervals by hypercubes or more general Borel sets. The modified coupling function \( \tilde{f} \) introduced before is replaced by \( \hat{f} \):

\[
\hat{f}(u^0; u^g, g \in G|w^{-1}_{-\infty}) = \begin{cases} 
\tilde{f}(u^0|w^{-1}_{-\infty}), & u^0 \geq a_1, \\
\tilde{f}(a_1 u^w|w^{-1}_{-\infty}), & u^0 < a_1, \ w_1 \in G.
\end{cases}
\]

which is not difficult to check that remains a coupling function for \( p \). As a corollary to Theorem 2 by collecting together the two previous remarks, we can construct a backward coalescence time (and thus a perfect simulation algorithm) for some interesting class of kernels \( p \). Notice that the last condition in the following corollary has the purpose of ensuring that \( a_0 = 0 \).

**Corollary 2.** Suppose that \( p : G \times G^{-N} \to [0,1] \) is a kernel on the finite state space \( G \). Define the oriented graph \( \mathcal{G} \) with set of vertices \( G \) and the set of arcs \( \mathcal{A} = \{(w,g) \in G^2 : a_1(g|w) > 0\} \). Suppose

1. \( \sum_{n=1}^{\infty} \prod_{m=1}^{n} a_m = \infty; \)
2. \( \mathcal{G} \) has a single irreducible class which is aperiodic;
3. for any \( g \in G \), there exists \( w \in G \) such that \( (w,g) \notin \mathcal{A} \).

Then it is possible to construct a backward coalescence time for the coupling function \( \hat{f} \).

**Example 3.** The previous result covers some generalized random walks on a finite directed graph \( \mathcal{G} = (G,A) \). Before defining this kind of processes, we define the set of one-sided infinite paths in \( \mathcal{G} \)

\[
\mathcal{C} = \{w^{-1}_{-\infty} \in G^{-N} : (w_{-(k+1)},w_{-k}) \in \mathcal{A}, k \in N_+ \}.
\]

Generalized random walks on \( \mathcal{G} = (G,A) \) are processes compatible with a kernel \( p \) over the alphabet \( G \) with the properties:

- if \( (g,w) \notin \mathcal{A} \) then, for all \( w^{-1}_{-\infty} \in G^{-N} \) with \( w_{-1} = w \), \( p(g|w^{-1}_{-\infty}) = 0 \);
- if \( (g,w) \in \mathcal{A} \), there exists \( \epsilon > 0 \) s.t. for all \( w^{-1}_{-\infty} \in \mathcal{C} \) with \( w_{-1} = w \), \( p(g|w^{-1}_{-\infty}) > \epsilon \).

The first property implies that \( \mathcal{H} \subset \mathcal{C} \), whereas the second ensures the opposite inclusion. Moreover

\[
a_1(g|w) = \inf \{p(g|w^{-1}_{-\infty}) : w^{-1}_{-\infty} \in \mathcal{C}, w_{-1} = w\} > \epsilon > 0,
\]

if \( (g,w) \in \mathcal{A} \) is an arc of \( \mathcal{G} \), otherwise it is clearly \( a_1(g|w) = 0 \). Thus we can get the set \( \mathcal{A} \) from the kernel \( p \) as indicated in Corollary 2. Therefore if the graph \( \mathcal{G} \) satisfies conditions (ii) and (iii) and the sequence \( \{a_k, k \in N_+\} \) satisfies condition (i), the previous Corollary allows to prove the existence and uniqueness of the generalized random walk, and the feasibility of a perfect simulation algorithm for sampling it.
The result of this section can be extended to cover the case \( a_1 = \cdots = a_l = 0, a_{l+1} > 0 \), for some \( l \geq 1 \). In this case the maximal coupling function depends on at least \( l + 1 \) variables hence it induces a markovian kernel \( M \) on the state space \( G^{l+1} \). The changes to the statement of Theorem 2 are rather straightforward.

References

[AG07] Søren Asmussen and Peter W. Glynn. *Stochastic simulation: algorithms and analysis*. Springer, New York, 2007.

[BB03] François Baccelli and Pierre Brémaud. *Elements of queueing theory*. Springer-Verlag, Berlin, 2003.

[Ber87] Henry Berbee. Chains with infinite connections: uniqueness and Markov representation. *Probab. Theory Related Fields*, 76(2):243–253, 1987.

[CFF02] Francis Comets, Roberto Fernández, and Pablo A. Ferrari. Processes with long memory: regenerative construction and perfect simulation. *Ann. Appl. Probab.*, 12(3):921–943, 2002.

[DSP08] Emilio De Santis and Mauro Piccioni. Exact simulation for discrete time spin systems and unilateral fields. *Methodol. Comput. Appl. Probab.*, 10(1):105–120, 2008.

[FFG02] Pablo A. Ferrari, Roberto Fernández, and Nancy L. García. Perfect simulation for interacting point processes, loss networks and Ising models. *Stochastic Process. Appl.*, 102(1):63–88, 2002.

[FT98] S. G. Foss and R. L. Tweedie. Perfect simulation and backward coupling. *Comm. Statist. Stochastic Models*, 14(1-2):187–203, 1998.

[Häg02] O. Häggström. *Finite Markov chains and algorithmic applications*. Cambridge University Press, Cambridge, 2002.

[HS00] Olle Häggström and Jeffrey E. Steif. Propp-Wilson algorithms and finitary codings for high noise Markov random fields. *Combin. Probab. Comput.*, 9(5):425–439, 2000.

[IG90] Marius Iosifescu and Şerban Grigorescu. *Dependence with complete connections and its applications*. Cambridge University Press, Cambridge, 1990.
[Lal00] S. P. Lalley. Regeneration in one-dimensional Gibbs states and chains with complete connections. Resenhas, 4(3):249–281, 2000.

[MG98] D. J. Murdoch and P. J. Green. Exact sampling from a continuous state space. Scand. J. Statist., 25(3):483–502, 1998.

[Møl01] Jesper Møller. A review of perfect simulation in stochastic geometry. In Selected Proceedings of the Symposium on Inference for Stochastic Processes (Athens, GA, 2000), pages 333–355. Inst. Math. Statist., Beachwood, OH, 2001.

[PW96] James Gary Propp and David Bruce Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. Random Structures Algorithms, 9(1-2):223–252, 1996.

[Wil91] David Williams. Probability with martingales. Cambridge University Press, Cambridge, 1991.