EXACT SAMPLING USING BRANCHING PARTICLE SIMULATION

Christophe Andrieu, Nicolas Chopin, Arnaud Doucet, Sylvain Rubenthaler

EXACT SAMPLING USING BRANCHING PARTICLE SIMULATION

CHRISTOPHE ANDRIEU, NICOLAS CHOPIN, ARNAUD DOUCET, AND SYLVAIN RUBENTHALER

Abstract. Particle methods, also known as Sequential Monte Carlo methods, are a popular set of computational tools used to sample approximately from non-standard probability distributions. A variety of convergence results ensure that, under weak assumptions, the distribution of the particles converges to the target probability distribution of interest as the number of particles increases to infinity. Unfortunately it can be difficult to determine practically how large this number needs to be to obtain a reliable approximation. We propose here a procedure which allows us to return exact samples. The proposed algorithm relies on the combination of an original branching variant of particle Markov chain Monte Carlo methods and dominated coupling from the past.

1. Introduction

We are given measurable spaces $(E_1, \mathcal{E}_1)$, $(E_2, \mathcal{E}_2)$, $\dots$, M_1 a probability measure on E_1, for each $k \geq 2$, a transition kernel M_k from E_{k-1} to E_k and bounded measurable potentials $(G_k)_{k\geq 1}$ ($G_k : E_k \to \mathbb{R}^+$, $\mathbb{R}^+$ equipped with the Lebesgue tribe). All densities and kernels are supposed to have a density with respect to some reference measures on E_k ($k = 1, 2, \dots, T$). Moreover, in the following, densities on enumerable sets will always be taken with respect to the counting measure. In the case we write a density on a space defined as a product of spaces E_i, the reference measure will be the product of the measures mentioned above. We want to draw samples according to the law (on paths of length T) defined for any measurable function f by

$$\pi(f) = \frac{\mathbb{E}\left(f(X_1, \dots, X_T)\prod_{k=1}^{T-1} G_k(X_k)\right)}{\mathbb{E}\left(\prod_{k=1}^{T-1} G_k(X_k)\right)} \tag{1.1}$$

where $(X_k)_{k\geq 1}$ is Markov with initial law M_1 and transitions $(M_k)_{k\geq 2}$ (for all $k \geq 1$, X_k takes values in E_k). For all $n \in \mathbb{N}^*$, we note $[n] = \{1, \dots, n\}$. We set $Z_T = \mathbb{E}\left(\prod_{i=1}^{T-1} G_i(X_i)\right)$. Then π has the following density at $(x_1, \dots, x_T) \in E_1 \times \cdots \times E_T$:

$$\pi(x_1, \dots, x_T) = \frac{1}{Z_T} M_1(x_1) \prod_{k=1}^{T-1} G_k(x_k) M_{k+1}(x_k, x_{k+1})\,. \tag{1.2}$$

In the following, all the random variables are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

We would like to draw a random variable according to the law π. Moreover, we want the complexity of the algorithm used to draw this variable to grow at most as a polynomial in T.

There is a straightforward way of sampling according to the law π. Suppose $\|G_k\|_\infty \leq 1$ for all k, for the sake of simplicity (the argument could be adapted to our hypotheses). We can use an accept-reject scheme by sampling the Markov chain $(X_1, X_2, \dots, X_T)$ and U of uniform law on $[0, 1]$ until $U \leq G_1(X_1) \times \dots G_{T-1}(X_{T-1})$. The first accepted $(X_1, \dots, X_T)$ is of law π. Unfortunately, the cost of this procedure is, in expectation, exponential in T.

In [ADH10], the authors propose a new algorithm called PMCMC (Particle Markov Chain Monte Carlo). Applied to our case, it is essentially a Markov chain in some space of genealogy of particles containing $E_1 \times E_2 \times \cdots \times E_T$ as a subspace. It happens that this Markov chain has a

Date: 18th October 2016.

Key words and phrases. Coupling from the past, Exact simulation, Particle methods, Particle Markov chain Monte Carlo Methods.

stationary law whose marginal in $E_1 \times \cdots \times E_T$ is exactly π. In order to generate genealogies, the PMCMC algorithm uses an SMC scheme (Sequential Monte-Carlo).

Given a Markov chain having this interesting stationary measure, we want to use CFTP (Coupling From The Past) to generate random variables which are exactly of law π. There are essentially two cases where CFTP is easy to apply: Markov chains in a partially ordered space (such as in the Ising's model) and Metropolis-Hastings Markov chain with independent proposals. We recall here a few references on CFTP: [FT98, PW96] (fundamental papers), [Ken05] (survey on perfect simulation), [CT02] (Metropolis-Hastings with independent proposals). In the present paper, we use a Metropolis-Hastings with independent proposals to build an algorithm which produces samples of law π. The proposals are inspired by the PMCMC algorithm. However, the use of the SMC makes the stopping time in the CFTP difficult to control. So we replace the SMC part by what we call "branching processes", which is another way of producing genealogies of particles.

The outline of the paper is the following. In Section 2, we define branching processes, which are random variable on complex spaces. In Section 3, we describe a Markov chain in the space of trajectories ($E_1 \times E_2 \times \cdots \times E_T$) having the invariant law π, this construction is inspired by the PMCMC algorithm. In Section 4,we use CFTP to build an algorithm producing samples of law π. In Section 5, we discuss the implementation and the complexity of this algorithm on two examples.

2. Densities of branching processes

2.1. **Branching process.** We first introduce some definitions concerning the elements of $(\mathbb{N}^*)^k$ for $k = 1, 2, \dots$. If $1 \le q \le n$ and $\boldsymbol{i} = (i_1, \dots, i_n) \in (\mathbb{N}^*)^n$, we define $\boldsymbol{i}(q) = (i_1, \dots, i_q)$ and we say that $\boldsymbol{i}(q)$ is an <u>ancestor</u> of $\boldsymbol{i}$ and we denote this relation by $\boldsymbol{i}(q) \prec \boldsymbol{i}$; we will also say that $\boldsymbol{i}$ is a <u>descendant</u> of $\boldsymbol{i}(q)$. We introduce the following notation: if $\boldsymbol{i} = (i_1, \dots, i_k) \in (\mathbb{N}^*)^k$ and $j \in \mathbb{N}^*$, $(\boldsymbol{i}, j) = (i_1, \dots, i_k, j) \in (\mathbb{N}^*)^{k+1}$. For $\boldsymbol{i}$, $\boldsymbol{j}$ in $(\mathbb{N}^*)^k$ ($k \in \mathbb{N}^*$), we write $\boldsymbol{i} \le \boldsymbol{j}$ if $\boldsymbol{i}$ is smaller that $\boldsymbol{j}$ in the alphanumerical sense. For example, in the case $k = 2$, $(1,1) \le (1,2) \le (1,3) \le \cdots \le (2,1) \le (2,2) \le \dots$.

We now define a branching process. We start with n_1 particles, i.i.d. with law M_1 (n_1 is a fixed number). We then proceed recursively through time to build a genealogical tree. At time 1, the particles are denoted by $X_1^1, \dots, X_1^{n_1}$. We set

$$S_1 = \{1, 2, \dots, n_1\}. \tag{2.1}$$

At time k, the particles are denoted by $(X_k^{\boldsymbol{i}})_{\boldsymbol{i} \in S_k}$, where S_k is a finite subset of $(\mathbb{N}^*)^k$, and the number of particles is $N_k = \# S_k$ (the cardinality of S_k). For $\boldsymbol{i} \in S_k$, we say that $X_k^{\boldsymbol{i}}$ is the <u>position of the particle</u> indexed by $\boldsymbol{i}$, or, in a shorter way, the position of $\boldsymbol{i}$. In an abuse of notation, in the case $\boldsymbol{j} \prec \boldsymbol{i}$ ($\boldsymbol{j} \in S_q$), we also say that $X_q^{\boldsymbol{j}}$ is an ancestor of $X_k^{\boldsymbol{i}}$ and that $X_k^{\boldsymbol{i}}$ is a descendant of $X_k^{\boldsymbol{j}}$. Starting from the particles at time $k \le T - 1$ with particles $(X_k^{\boldsymbol{i}})_{\boldsymbol{i} \in S_k}$ ($S_k \subset (\mathbb{N}^*)^k$), the system evolves in the following manner:

- For each $\boldsymbol{i} \in S_k$, the number of children of $\boldsymbol{i}$ is a random variable $N_{k+1}^{\boldsymbol{i}}$ with law f_{k+1} such that :
$$\mathbb{P}(N_{k+1}^{\boldsymbol{i}} = j \,|\, X_k^{\boldsymbol{i}}) = f_{k+1}(G_k(X_k^{\boldsymbol{i}}), j).$$
Here, f_{k+1} is a law with a parameter $G_k(X_k^{\boldsymbol{i}})$, we will define this law precisely later. We suppose that
$$f_{k+1}(G_k(x), 0) = 1 - \alpha_{k+1} G_k(x) \tag{2.2}$$
for some α_{k+1} in $[0, 1/\|G_k\|_\infty]$. This will remain true trough all the paper. The variables $N_{k+1}^{\boldsymbol{i}}$ ($\boldsymbol{i} \in S_k$) are independent. The total number of particles at time $k+1$ is then $N_{k+1} = \sum_{\boldsymbol{i} \in S_k} N_k^{\boldsymbol{i}}$.
- If $N_{k+1} \ne 0$, we draw σ_{k+1} uniformly in $\mathfrak{S}_{N_{k+1}}$ (the N_{k+1}-th symmetric group). If $N_{k+1} = 0$, we use the convention $\mathfrak{S}_{N_{k+1}} = \emptyset$ and the system stops here.

- We order S_k alphanumerically: $S_k = \{\boldsymbol{i}_1, \dots, \boldsymbol{i}_{N_k}$ with $\boldsymbol{i}_1 \le \dots \le \boldsymbol{i}_{N_k}\}$. For $r \in [N_k]$, we set
$$C_{k+1}^r = \left\{ 1 + \sum_{l=1}^{r-1} N_{k+1}^{\boldsymbol{i}_l}, 2 + \sum_{l=1}^{r-1} N_{k+1}^{\boldsymbol{i}_l}, \dots, \sum_{l=1}^{r} N_{k+1}^{\boldsymbol{i}_l} \right\} .$$
We set
$$S_{k+1} = \cup_{r=1}^{N_k} \cup_{j \in \sigma_{k+1}(C_{k+1}^r)} (\boldsymbol{i}_r, j) .$$
For $r \in [N_k]$, $j \in \sigma_{k+1}(C_{k+1}^r)$, we draw $X_{k+1}^{(\boldsymbol{i}_r,j)} \sim M_{k+1}(X_k^{\boldsymbol{i}_r}, .)$ in E_{k+1}. To simplify the notation, we write $C(\boldsymbol{i}_r) = C_{k+1}^r$ for all $r \in [N_k]$ and $C(\boldsymbol{i}) = C(\boldsymbol{i}_r)$ if $\boldsymbol{i} = \boldsymbol{i}_r \in S_k$. We can then write that, for $\boldsymbol{i} \in S_k$, the descendants of $\boldsymbol{i}$ at time $k+1$ are the $(\boldsymbol{i}, j)$ for $j \in \sigma_{k+1}(C(\boldsymbol{i}))$.

Such a system takes values in the space
$$E = \{(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k}) : S_k \in (\mathbb{N}^*)^k \, \forall k \in [T] , \text{ satisfying condition } \mathrm{H}_T\} ,$$
where the conditions H_1, H_2, ... are defined recursively:

- H_1 is: $S_1 = [n_1]$, $\forall \boldsymbol{i} \in S_1$, $x_1^{\boldsymbol{i}} \in E_1$,
- $(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k})$ satisfies condition H_T if
  - $S_T \subset (\mathbb{N}^*)^T$
  - $(S_1, \dots, S_{T-1}, (x_k^{\boldsymbol{i}})_{k \in [T-1], \boldsymbol{i} \in S_k})$ satisfies condition H_{T-1},
  - $\forall \boldsymbol{i} \in S_T$, $\exists \boldsymbol{j} \in S_{T-1}$ such that $\boldsymbol{j} \prec \boldsymbol{i}$

For $(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k}))$ in E, we set
$$\forall k \in [T-1] , \, \forall \boldsymbol{i} \in S_k , \, n_{k+1}^{\boldsymbol{i}} = \#\{\boldsymbol{j} \in S_{k+1} \, : \, \boldsymbol{i} \prec \boldsymbol{j}\} , \, n_{k+1} = \sum_{\boldsymbol{i} \in S_k} n_{k+1}^{\boldsymbol{i}} .$$
We remark that for any element $(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k})$ of E and $k \in \{1, 3, \dots, T-1\}$, if $n_{k+1} = 0$ then $n_{k+1} = n_{k+2} = \dots = n_T = 0$, $S_{k+1} = S_{k+2} = \dots = S_T = \emptyset$. A random variable on the space E will be called a branching process.

At a point $(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k})$ in E, the density of the branching process defined above is given by

$$q(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k}) = \prod_{\boldsymbol{i} \in S_1} M_1(x_1^{\boldsymbol{i}}) \prod_{k=1}^{T-1} \left\{ \prod_{\boldsymbol{i} \in S_k} f_{k+1}(G_k(x_k^{\boldsymbol{i}}), n_{k+1}^{\boldsymbol{i}}) \times \frac{1}{n_{k+1}!} \times \prod_{\boldsymbol{i} \in S_k} \prod_{\boldsymbol{j} \in S_{k+1} : \boldsymbol{i} \prec \boldsymbol{j}} M_{k+1}(x_k^{\boldsymbol{i}}, x_{k+1}^{\boldsymbol{j}}) \right\} . \tag{2.3}$$

In the following, we use the short cut notation $(S_{1:T}, (x_k^{\boldsymbol{i}}))$ for a point in E. For $\chi \in E$, we write $S_k(\chi)$, $k \in [T]$ for the corresponding subsets of $(\mathbb{N}^*)^k$, $k \in [T]$, we write $N_k(\chi)$ for the number of particles at each time step $k \in [T]$. We will use the same notations for a point $\chi \in E'$ (E' defined below).

2.2. **Proposal distribution.** We introduce the space
$$E' = E \times (\mathbb{N}^*)^T .$$
Suppose we draw a branching process $(S_{1:T}, (X_k^{\boldsymbol{i}}))$ with law q and then draw $\boldsymbol{B}$ uniformly in S_T if $N_T \ge 1$, and set $\boldsymbol{B} = (1, 1, \dots, 1)$ if $N_T = 0$. This random variable takes values in E' and, at a point $(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k}, \boldsymbol{b})$ with $(S_{1:T}, (x_k^{\boldsymbol{i}})) \in E$ such that $n_T \ge 1$, $\boldsymbol{b} \in S_T$, it has the density

$$\widehat{q}(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k}, \boldsymbol{b}) = q(S_1, \dots, S_T, (x_k^{\boldsymbol{i}})_{k \in [T], \boldsymbol{i} \in S_k}) \times \frac{1}{n_T} . \tag{2.4}$$

We call this the proposal distribution. One can view the space E' as the space of branching processes where a particular trajectory is singled out, we will call it the coloured trajectory. At a point $(S_{1:T}, (x_k^{\boldsymbol{i}}), \boldsymbol{b})$, what we call the coloured trajectory is $(x_1^{\boldsymbol{b}(1)}, \dots, x_T^{\boldsymbol{b}(T)})$ in the case $n_T \ge 1$; in the case $n_T = 0$, the coloured trajectory is $(x_1^1, x_2^{(1,1)}, \dots, x_q^{(1,\dots,1)})$ where $q = \max\{k \, : \, n_k \ge 1\}$.

In Figure 2.1, we have a representation of a the realisation variable of law $\widehat{q}$ on the left. The sets E_1, E_2, ... are equal to $\mathbb{R}$, we draw $n_1 = 3$ points at time 1 and a full branching process according to the law q. The terminal time T is equal to 4. The number of particles at time T is $N_T = 7$. Two particles are linked when one is the children of the other. One particle is chosen uniformly amongst the particles at time T. This particle is coloured and its ancestors are coloured. This makes the coloured trajectory.

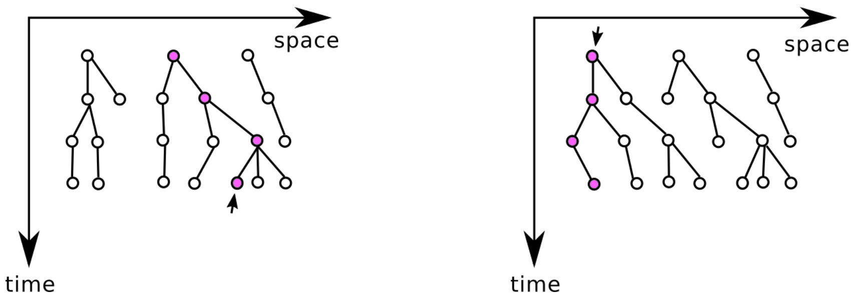

FIGURE 2.1. Examples

2.3. **Target distribution.** Let us denote by $\widehat{f}(g,.)$ the law $f(g,.)$ conditioned to be ≥ 1, that is: for all $g \geq 0$, $i \geq 1$, $k \geq 1$,

$$\widehat{f}_k(g,i) = \frac{f_k(g,i)}{1 - f_k(g,0)}\,. \tag{2.5}$$

This quantity is not defined in the case $g = 0$ but we will not need it in this case. An alternative way of building a branching process with a coloured trajectory is to draw a trajectory with law π, say that is is the coloured trajectory and then build a branching process conditioned to contain this trajectory. The indexes of the coloured trajectory embedded in the branching process are denoted by a random variable $\boldsymbol{B}$. The first coordinate $\boldsymbol{B}(1)$ is chosen uniformly in $[n_1]$. The other coordinates are deduced from the branching process in the following way: suppose that, at time $k+1$, the random permutation of the branching process is σ_{k+1} and the numbers of children are $(N^{\boldsymbol{i}}_{k+1})_{\boldsymbol{i}\in S_k}$, we set $\boldsymbol{B}_{k+1} = \sigma_{k+1}(\min\{C(\boldsymbol{B}(k))\})$ and $\boldsymbol{B}(k+1) = (\boldsymbol{B}(k), \boldsymbol{B}_{k+1})$. We thus introduce what we call the target distribution. Its support is contained in $\{(S_1, \dots, S_T, (x^{\boldsymbol{i}}_k)_{k\in[T],\boldsymbol{i}\in S_k}, \boldsymbol{b}) \in E' : n_T \geq 1\}$ and it has the density:

$$\begin{aligned}
&\widehat{\pi}(S_1, \dots, S_T, (x^{\boldsymbol{i}}_k)_{k\in[T],\boldsymbol{i}\in S_k}, \boldsymbol{b}) \\
&\qquad = \frac{1}{n_1}\pi\left(x_1^{\boldsymbol{b}(1)}, ..., x_T^{\boldsymbol{b}(T)}\right) \prod_{\boldsymbol{i}\in S_1\,,\,\boldsymbol{i}\neq\boldsymbol{b}(1)} M_1(x_1^{\boldsymbol{i}}) \\
&\times \prod_{k=1}^{T-1}\left\{\left(\prod_{\boldsymbol{i}\in S_k,\boldsymbol{i}\neq\boldsymbol{b}(k)} f_{k+1}(G_k(x^{\boldsymbol{i}}_k), n^{\boldsymbol{i}}_{k+1})\right) \times \widehat{f}_{k+1}(G_k(x_k^{\boldsymbol{b}(k)}), n_{k+1}^{\boldsymbol{b}(k)}) \times \frac{1}{n_{k+1}!}\right. \\
&\qquad\qquad \left. \times \left(\prod_{\boldsymbol{i}\in S_k,\boldsymbol{i}\neq\boldsymbol{b}(k)}\ \prod_{\boldsymbol{j}\in S_{k+1}\setminus\{\boldsymbol{b}(k+1)\},\boldsymbol{i}\prec\boldsymbol{j}} M_{k+1}(x^{\boldsymbol{i}}_k, x^{\boldsymbol{j}}_{k+1})\right)\right\},
\end{aligned} \tag{2.6}$$

where the term $\widehat{f}_{k+1}(G_k(x_k^{\boldsymbol{b}(k)}), n_{k+1}^{\boldsymbol{b}(k)})$ corresponds to the simulation of the number of offsprings of the particle $x_k^{\boldsymbol{b}(k)}$. Using (1.2). (2.2) and (2.5), we can rewrite $\widehat{\pi}$ into

$$\begin{aligned}
&\widehat{\pi}(S_1, \dots, S_T, (x^{\boldsymbol{i}}_k)_{k\in[T],\boldsymbol{i}\in S_k}, \boldsymbol{b}) \\
&= \frac{1}{n_1} \times \prod_{i=1}^{n_1} M_1(x_1^i) \times \prod_{k=1}^{T-1}\left\{\frac{1}{n_k!}\prod_{i\in S_k}\left(f_{k+1}(G_k(x^{\boldsymbol{i}}_k), n^{\boldsymbol{i}}_{k+1}) \times \prod_{j\in S_{k+1},\boldsymbol{i}\prec\boldsymbol{j}} M_{k+1}(x^{\boldsymbol{i}}_k, x^{(\boldsymbol{i},j)}_{k+1})\right)\right\}
\end{aligned} \tag{2.7}$$

$$\times \frac{1}{Z_T \alpha_1 \dots \alpha_{T-1}} .$$

In Figure 2.1, we have a representation of a the realisation variable of law $\widehat{\pi}$ on the right. We draw a trajectory with law π. This is the coloured trajectory. We then draw a forest of law q conditioned to contain this coloured trajectory. The number of particles at time T is $\overline{N}_T = 7$.

2.4. **Ratio of densities.** We deduce from (2.3), (2.4) and (2.7), that at a point in the support of $\widehat{\pi}$, the ratio of $\widehat{\pi}$ and $\widehat{q}$ is equal to

$$\frac{\widehat{\pi}((x_k^{\boldsymbol{i}})_{k\in[T],\boldsymbol{i}\in S_k},(n_k^{\boldsymbol{i}})_{k\in\{2,\dots,T\},\boldsymbol{i}\in S_{k-1}},(\sigma_k)_{k\in\{2,\dots,T\}},\boldsymbol{b})}{\widehat{q}((x_k^{\boldsymbol{i}})_{k\in[T],\boldsymbol{i}\in S_k},(n_k^{\boldsymbol{i}})_{k\in\{2,\dots,T\},\boldsymbol{i}\in S_{k-1}},(\sigma_k)_{k\in\{2,\dots,T\}},\boldsymbol{b})} = \frac{n_T}{n_1 Z_T \alpha_1 \dots \alpha_{T-1}} . \tag{2.8}$$

3. A Markov chain on $E_1 \times \cdots \times E_T$

We now define a Markov kernel Q on $E_1 \times \cdots \times E_T$. We start from a path $(x_1, \dots, x_T) \in E_1 \times \cdots \times E_T$. We will move to a new path in several steps.

(1) Draw of a conditional forest. We sample a random variable χ with law $\widehat{\pi}$ conditionally on $(X_1^{\boldsymbol{B}(1)}, \dots, X_T^{\boldsymbol{B}(T)}) = (x_1, \dots, x_T)$. We use for this the expression (2.6). Such a sampling can be done recursively in $k \in [T]$ in the following way.
   - We take $\boldsymbol{B}(1) = \boldsymbol{B}_1$ uniformly in $[n_1]$. We take $(X_1^{\boldsymbol{i}})_{\boldsymbol{i}\in S,\boldsymbol{i}\neq \boldsymbol{B}_1}$ i.i.d. with law M_1.
   - Suppose we have sampled $((X_q^{\boldsymbol{i}})_{1\le q\le k,\boldsymbol{i}\in S_q}, (N_q^{\boldsymbol{i}})_{2\le q\le k,\boldsymbol{i}\in S_{q-1}}, S_k)$ for $k \le T-1$. For $\boldsymbol{i} \in S_k$, we take $N_{k+1}^{\boldsymbol{i}}$ with law $f_{k+1}(G_k(X_k^{\boldsymbol{i}}), .)$ if $\boldsymbol{i} \neq \boldsymbol{B}(k)$ and $N_{k+1}^{\boldsymbol{B}(k)}$ with law $\widehat{f}_{k+1}(G_k(X_k^{\boldsymbol{B}(k)}), .)$. We set $N_{k+1} = \sum_{\boldsymbol{i}\in S_k} N_k^{\boldsymbol{i}}$. We draw σ_{k+1} uniformly in $\mathfrak{S}_{N_{k+1}}$. We set $\boldsymbol{B}_{k+1} = \sigma_{k+1}(\min\{C(\boldsymbol{B}(k))\})$, $\boldsymbol{B}(k+1) = (\boldsymbol{B}(k), \boldsymbol{B}_{k+1})$. We set $S_{k+1} = \cup_{\boldsymbol{i}\in S_k} \cup_{j\in\sigma_{k+1}(C(\boldsymbol{i}))} (\boldsymbol{i}, j)$.
     - For $\boldsymbol{i} \in S_k$, $\boldsymbol{i} \neq \boldsymbol{B}(k)$, $j \in \sigma_{k+1}(C(\boldsymbol{i}))$, we take $X_{k+1}^{(\boldsymbol{i},j)}$ with law $M_{k+1}(X_k^{\boldsymbol{i}}, .)$.
     - For $\boldsymbol{i} = \boldsymbol{B}(k)$, $j \in \sigma_{k+1}(C(\boldsymbol{i}))$, $j \neq \boldsymbol{B}_{k+1}$, we take $X_{k+1}^{(\boldsymbol{i},j)}$ with law $M_{k+1}(X_k^{\boldsymbol{i}}, .)$.

(2) Proposal. We draw a proposal $\overline{\chi}$ with law $\widehat{q}$. It contains a coloured trajectory $(\overline{X}_1^{\overline{\boldsymbol{B}}(1)}, \dots, \overline{X}_T^{\overline{\boldsymbol{B}}(T)})$.

(3) Accept/reject step. We move to $(\overline{X}_1^{\overline{\boldsymbol{B}}(1)}, \dots, \overline{X}_T^{\overline{\boldsymbol{B}}(T)})$ with probability $\min\left(1, \frac{N_T(\overline{\chi})}{N_T(\chi)}\right)$ and we stay in $(x_1, \dots, x_T)$ with probability

$$1 - \min\left(1, \frac{N_T(\overline{\chi})}{N_T(\chi)}\right) . \tag{3.1}$$

Theorem 3.1. *The law π is invariant for the kernel Q.*

Proof. Suppose we start with a random variable $(X_1, \dots, X_T)$ with law π. Going through step 1 of the construction above, we get a random variable

$$\chi = ((X_k^{\boldsymbol{i}})_{k\in[T],\boldsymbol{i}\in S_k}, (N_k^{\boldsymbol{i}})_{k\in\{2,\dots,T\},\boldsymbol{i}\in S_{k-1}}, (\sigma_k)_{k\in\{2,\dots,T\}}, \boldsymbol{B})$$

in E' such that $(X_1^{\boldsymbol{B}(1)}, \dots, X_T^{\boldsymbol{B}(T)}) = (X_1, \dots, X_T)$. By (2.6), χ has the law $\widehat{\pi}$. We draw a random variable $\overline{\chi}$ with law $\widehat{q}$ as in step 2 above. We then proceed to the step 3 above. Let U be a uniform variable in $[0, 1]$. We set

$$\widehat{\chi} = \begin{cases} \overline{\chi} & \text{if } U \le \min\left(1, \frac{N_T(\overline{\chi})}{N_T(\chi)}\right) , \\ \chi & \text{otherwise.} \end{cases}$$

The result of the random move by the Markov kernel Q is the coloured trajectory of $\widehat{\chi}$. By (2.8), we have that

$$\frac{N_T(\overline{\chi})}{N_T(\chi)} = \frac{\widehat{\pi}(\overline{\chi})\widehat{q}(\chi)}{\widehat{q}(\overline{\chi})\widehat{\pi}(\chi)} , \tag{3.2}$$

and so $\widehat{\chi}$ is of law $\widehat{\pi}$ (we recognise here the accept-reject step of a Metropolis-Hastings algorithm). This finishes the proof. $\square$

Remark 3.2. We use permutations in the definitions of the proposal and target distribution. It might look as an unnecessary complication. However, this is what makes that $\widehat{\pi}$ and $\widehat{q}$ have almost the same support and that the acceptance ratio on the right-hand side of Equation (3.2) above simplifies nicely into $N_T(\overline{\chi})/N_T(\chi)$.

4. Algorithms

4.1. **Simulation of a branching process.** One might be worried whether a process of law q might be such that N_P is very big or equal to 0 with a high probability. Such events are undesirable in our simulations, as it will be seen later. For all $k \geq 1$, $g \in [0; \|G_k\|_\infty]$, we want

$$f_{k+1}(g,0) = 1 - \alpha_{k+1} g$$

(for some parameter $\alpha_{k+1} \in (0, 1/\|G_k\|_\infty])$. Suppose we take the following simple law:

$$f_{k+1}(g,i) = \alpha_{k+1} p_{k+1,i} g \text{ for } i \in [q_{k+1}] \tag{4.1}$$

and $f_{k+1}(g,i) = 0$ for $i \geq q_{k+1}+1$, with $p_{k+1,i} \geq 0$ for $i \in [q_{k+1}]$, and with some integer q_{k+1} to be chosen. For example, one could choose $p_{k+1,i} = 1/q_{k+1}$ for all $i \in [q_{k+1}]$, in which case:

$$f_{k+1}(g,i) = g \times \frac{\alpha_{k+1} i}{q_{k+1}}\,. \tag{4.2}$$

Suppose we have built a branching process up to time k. We define a measure π_k on E_k by its action on test functions:

$$\pi_k(f) = \frac{\mathbb{E}(f(X_k)\prod_{i=1}^{k-1} G_i(X_i))}{\mathbb{E}(\prod_{i=1}^{k-1} G_i(X_i))}\,, \tag{4.3}$$

where $(X_1, \dots, X_T)$ is a non-homogeneous Markov chain with initial law M_1 and transitions M_2, M_3, ... M_T. Suppose we make a simulation of a branching process up to time k. The particles at time k are denoted by $(X_k^{\boldsymbol{i}})_{\boldsymbol{i}\in S_k}$. The empirical measure

$$\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k} \delta_{X_k^{\boldsymbol{i}}}$$

is such that, for all bounded measurable $\varphi : E_k \to \mathbb{R}^+$,

$$\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k} \varphi(X_k^{\boldsymbol{i}}) \xrightarrow[n_1\to+\infty]{\mathbb{P}} \pi_k(\varphi)$$

and $N_k \xrightarrow[n_1\to+\infty]{\mathbb{P}} +\infty$ (see Section 6.1 for the details). Knowing N_k, we want the expected number of children of the $(k+1)-$th generation to be N_k, that is:

$$\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}\sum_{j=1}^{q_{k+1}} j\alpha_{k+1}p_{k+1,j}G_k(X_k^{\boldsymbol{i}}) = \frac{1}{N_k}\sum_{i=1}^{N_k} m_{1,k+1}G_k(X_k^{\boldsymbol{i}}) = 1\,, \tag{4.4}$$

where

$$m_{1,k+1} = \sum_{j=1}^{q_{k+1}} \alpha_{k+1}p_{k+1,j}j\,. \tag{4.5}$$

This is true asymptotically when $n_1 \to +\infty$ if

$$m_{1,k+1}\pi_k(G_k) = 1\,. \tag{4.6}$$

Suppose now we have approximated π_k by an empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{Y_k^i}$ with some particles $(Y_k^i)_{1\leq i\leq N}$ coming from a SMC scheme (with N particles). We then have, for all bounded measurable φ,

$$\frac{1}{N}\sum_{i=1}^{N} \varphi(Y_k^i) \xrightarrow[N\to+\infty]{\mathbb{P}} \pi_k(\varphi)$$

We set

$$\pi_k^N(G_k) = \frac{1}{N}\sum_{i=1}^N G_k(Y_k^i)\,.$$

We suppose now, we are in the sub case of Equation (4.2). We take α_{k+1} such that

$$\frac{2}{\alpha_{k+1}\pi_k^N(G_k)} = \left\lceil \frac{2\|G\|_\infty}{\pi_k^N(G_k)} \right\rceil \text{ (upper integer part)}, \tag{4.7}$$

and

$$q_{k+1} = \left\lceil \frac{2\|G\|_\infty}{\pi_k^N(G_k)} \right\rceil - 1\,. \tag{4.8}$$

Then Equation (4.4) is true asymptotically when $n_1 \to +\infty$, $N \to +\infty$. With this choice of α_{k+1}, q_{k+1}, we have

$$q_{k+1} \geq 1\,,\ \alpha_{k+1}\|G_k\|_\infty \in \left[\frac{2}{3}, 1\right]\,. \tag{4.9}$$

We have here described a way of choosing the integers $(q_k)_{k\in\{2,\dots,T\}}$ before making a simulation of the branching process. The arguments given are purely heuristic. Other ways of calibrating $(q_k)_{k\in\{2,\dots,T\}}$ are possible, the only important thing is that these integers should be fixed before running the exact simulation algorithm described below.

We need that in practice, the number of particles in the branching process remains stable in time. The next Lemma tells us that if we can perfectly tune the parameters, the number of particles remain stable.

Lemma 4.1. *For the branching process described above, if $(\alpha_{k+1}, q_{k+1}, (p_{k+1,i})_{i\in[q_{k+1}]})_{k\in[T-1]}$ satisfy Equation (4.6) for all k and if the law f_{k+1} is described by Equation (4.1), then*

$$\mathbb{E}(N_k) = n_1\,,$$

for all k in $[T]$.

Proof. As in Section 6.1, we use the notation

$$m(x_1,\dots,x_N) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$$

for empirical measures. We want to prove the result by recurrence on k.

It is true for $k = 1$.

Let us suppose it is true in $k \geq 1$. We have,

$$\begin{aligned}
\mathbb{E}(N_{k+1}) &= \mathbb{E}\left(\sum_{\boldsymbol{i}\in S_k} N_{k+1}^{\boldsymbol{i}}\right) \\
&= \mathbb{E}\left(N_k m(X_k^{\boldsymbol{i}}, \boldsymbol{i}\in S_k)(G_k) m_{1,k+1}\right)\,.
\end{aligned}$$

And, by Equation (4.6),

$$\begin{aligned}
\mathbb{E}\left(N_k m(X_k^{\boldsymbol{i}}, \boldsymbol{i}\in S_k)(G_k) m_{1,k+1}\right) - \mathbb{E}(N_k) &= \mathbb{E}((N_k - n_1)(m(X_k^{\boldsymbol{i}}, \boldsymbol{i}\in S_k)(G_k) - \pi_k(G_k)) m_{1,k+1}) \\
&\quad + n_1\mathbb{E}((m(X_k^{\boldsymbol{i}}, \boldsymbol{i}\in S_k)(G_k) - \pi_k(G_k)) m_{1,k+1})\,.
\end{aligned}$$

By Lemma 6.2,

$$\mathbb{E}((m(X_k^{\boldsymbol{i}}, \boldsymbol{i}\in S_k)(G_k) - \pi_k(G_k)) m_{1,k+1}) \underset{n_1\to+\infty}{\longrightarrow} 0\,.$$

And we have

$$\begin{aligned}
&\frac{1}{n_1}\mathbb{E}((N_k - n_1)(m(X_k^{\boldsymbol{i}}, \boldsymbol{i}\in S_k)(G_k) - \pi_k(G_k)) m_{1,k+1}) \\
&\qquad \leq \frac{1}{n_1}\mathbb{E}((N_k - n_1)^2)^{1/2}\mathbb{E}((m(X_k^{\boldsymbol{i}}, \boldsymbol{i}\in S_k)(G_k) - \pi_k(G_k))^2 m_{1,k+1}^2)^{1/2}\,,
\end{aligned}$$

where

$$\mathbb{E}((m(X_k^{\boldsymbol{i}},\boldsymbol{i}\in S_k)(G_k)-\pi_k(G_k))^2 m_{1,k+1}^2)^{1/2} \underset{n_1\to+\infty}{\longrightarrow} 0$$

by Lemma 6.2, and

$$\begin{aligned}\frac{1}{n_1}\mathbb{E}((N_k-n_1)^2)^{1/2} &= \frac{1}{n_1}\mathbb{E}((\sum_{\boldsymbol{i}\in S_1}\#\{\boldsymbol{j}\in S_k\,:\,\boldsymbol{i}\prec\boldsymbol{j}\}-1)^2)\\ \text{(see Remark 4.3)} &= \mathbb{E}((\#\{\boldsymbol{j}\in S_k\,:\,1\prec\boldsymbol{j}\}-1)^2)\,,\end{aligned}$$

which does not depend on n_1. So we get

$$\frac{\mathbb{E}(N_{k+1})}{n_1}=1\,.$$

□

We set, for all $k\in[T]$,

$$\sigma_1(k)^2=\mathbb{V}(\#\{\boldsymbol{i}\in S_k\,:\,1\prec\boldsymbol{i}\})\,.$$

We have for all k in $[T]$,

$$\sigma(k)_1^2=\frac{\mathbb{V}\,(N_k)}{n_1}$$

(this does not depend on n_1, see Remark 4.3). We set for all $k\in[T-1]$,

$$m_{2,k+1}=\sum_{j=1}^{q_{k+1}}\alpha_{k+1}p_{k+1,j}j^2\,, \tag{4.10}$$

(we recall that $m_{1,k+1}$ is defined in Equation (4.5)).

Lemma 4.2. *Under the same assumptions as in Lemma 4.1, we have the recurrence relation, for all $k\in[T-1]$*

$$\sigma_1(k+1)^2=\sigma_1(k)^2+\pi_k(G_k)m_{2,k+1}+\pi_k(G_k^2)m_{1,k+1}^2-1\,.$$

Proof. We have $\sigma_1^2=0$. And for all $k\in[T-1]$,

$$\begin{aligned}\mathbb{E}(N_{k+1}^2) &= \mathbb{E}\left(\left(\sum_{\boldsymbol{i}\in S_k}N_{k+1}^{\boldsymbol{i}}\right)^2\right)\\ &= \mathbb{E}\left(\sum_{\boldsymbol{i}\in S_k}(N_{k+1}^{\boldsymbol{i}})^2+\sum_{\substack{\boldsymbol{i},\boldsymbol{j}\in S_k\\ \boldsymbol{i}\neq\boldsymbol{j}}}N_{k+1}^{\boldsymbol{i}}N_{k+1}^{\boldsymbol{j}}\right)\\ &= \mathbb{E}\left(\sum_{\boldsymbol{i}\in S_k}G_k(X_k^{\boldsymbol{i}})m_{2,k+1}+\sum_{\substack{\boldsymbol{i},\boldsymbol{j}\in S_k\\ \boldsymbol{i}\neq\boldsymbol{j}}}G_k(X_k^{\boldsymbol{i}})G_k(X_k^{\boldsymbol{j}})m_{1,k+1}^2\right)\,.\end{aligned}$$

We have (by Lemma 6.3)

$$\mathbb{E}\left(\sum_{\boldsymbol{i}\in S_k}G_k(X_k^{\boldsymbol{i}})m_{2,k+1}\right)=\mathbb{E}\left(n_1\sum_{1\prec\boldsymbol{i}\in S_k}G_k(X_k^{\boldsymbol{i}})m_{2,k+1}\right)=n_1\pi_k(G_k)m_{2,k+1}\,,$$

$$\mathbb{E}\left(\sum_{\substack{\boldsymbol{i},\boldsymbol{j}\in S_k\\ \boldsymbol{i}\neq\boldsymbol{j}}} G_k(X_k^{\boldsymbol{i}})G_k(X_k^{\boldsymbol{j}})m_{1,k+1}^2\right)-n_1^2=$$

$$\mathbb{E}\left(\sum_{\substack{\boldsymbol{i},\boldsymbol{j}\in S_k\\ \boldsymbol{i}\neq\boldsymbol{j}}} (G_k(X_k^{\boldsymbol{i}})G_k(X_k^{\boldsymbol{j}})m_{1,k+1}^2-1)\right)+\mathbb{E}(N_k(N_k-1))-n_1^2=$$

$$(\ N_k\text{takes its values in } \mathbb{N}\ )$$

$$n_1(n_1-1)\mathbb{E}\left(\sum_{1\prec\boldsymbol{i}\in S_k}\sum_{2\prec\boldsymbol{j}\in S_k}(G_k(X_k^{\boldsymbol{i}})G_k(X_k^{\boldsymbol{j}})m_{1,k+1}^2-1)\right)$$

$$+n_1\mathbb{E}\left(\sum_{\substack{1\prec\boldsymbol{i},\boldsymbol{j}\in S_k\\ \boldsymbol{i}\neq\boldsymbol{j}}}(G_k(X_k^{\boldsymbol{i}})^2m_{1,k+1}^2-1)\right)+n_1\sigma_k^2=$$

$$\text{(Lemma 6.3 and Remark 4.3)}$$

$$0+n_1(\pi_k(G_k^2)m_{1,k+1}^2-1)+n_1\sigma_k^2\,.$$

So

$$\sigma_1(k+1)^2=\sigma_1(k)^2+\pi_k(G_k)m_{2,k+1}+\pi_k(G_k^2)m_{1,k+1}^2-1\,.$$

□

4.2. **Representation of the Markov transition.** Suppose we have a trajectory $(x_1,\dots,x_T)\in E_1\times\cdots\times E_T$. We want to sample a random variable of law $Q((x_1,\dots,x_T),.)$. In practice, we do not have to make a simulation of the random permutations appearing in the Markov transition Q because the indexes are not used in the computation of the acceptance ratio in Equation (3.1). We can simply run a simulation of the positions of the particles and forget about their indexes. Having said this, we change the way we index the particles.

We take functions $(m_k)_{1\leq k\leq T}$, $(\varphi_k)_{k\leq 2\leq T}$, $(\widehat{\varphi}_k)_{2\leq k\leq T}$ such that $m_1:[0,1]\to E_1$ for all $k\in\{2,\dots,T\}$, $n\in\mathbb{N}^*$, $m_k:E_{k-1}\times[0,1]\to E_k$, $\varphi_k:\mathbb{R}_+\times[0,1]\to\mathbb{N}$, $\widehat{\varphi}_k:\mathbb{R}_+\times[0,1]\to\mathbb{N}$ and for a random variable U of uniform law on $[0,1]$,

$$m_1(U)\text{ is of law } M_1\,,$$

and for any $k\in[T-1]$, $x\in E_k$, $g\in\mathbb{R}_+$, $j\in\mathbb{N}$,

$$m_{k+1}(x,U)\text{ is of law } M_{k+1}(x,.)\,, \tag{4.11}$$

$$\mathbb{P}(\varphi_{k+1}(G_k(x),U)=j)=f_{k+1}(G_k(x),j)\,,$$

$$\mathbb{P}(\widehat{\varphi}_{k+1}(G_k(x),U)=j)=\widehat{f}_{k+1}(G_k(x),i)\,.$$

For $k\in[T-1]$, $x\in E_k$, $u\in[0,1]\mapsto\varphi_{k+1}(G_k(x),u)$ is the pseudo-inverse of the cumulative distribution function of the random variable of law $f_{k+1}(G_{k+1}(x),.)$, and $u\in[0,1]\mapsto\widehat{\varphi}_{k+1}(G_k(x),u)$ is the pseudo-inverse of the cumulative distribution function of the random variable of law $\widehat{f}_{k+1}(G_{k+1}(x),.)$.

Suppose now we are given a random variable $\Theta=(U_{\boldsymbol{i}},U'_{\boldsymbol{i}},V_{\boldsymbol{i}},V'_{\boldsymbol{i}},W_1,W_2)_{k\in\mathbb{N}^*,\boldsymbol{i}\in(\mathbb{N}^*)^k}$ made of a family of i.i.d. random variables of uniform law on $[0,1]$. We denote by $\mathcal{O}$ the space in which Θ takes its value. The space $\mathcal{O}$ is in bijection with $[0,1]^{\mathbb{N}}$. The set $[0,1]^{\mathbb{N}}$ is equipped with the

Cartesian product of the Lebesgue tribe on $[0,1]$ and so this bijection induces a tribe on $\mathcal{O}$. Using these variables and the functions above, we can build a random variable of law $Q((x_1,\dots,x_T),.)$. We start with a recursive construction of the conditional forest.

- We set $\boldsymbol{B}=(1,1,\dots,1)\in\mathbb{N}^T$. We set $X_1^{\boldsymbol{B}(1)}=x_1$. We set $X_1^i=m_1(U_i)$ for $i\in\{2,\dots,n_1\}$. We set $S_1=[n_1]$.
- Suppose we have $S_k\subset(\mathbb{N}^*)^k$ of cardinality N_k , containing $\boldsymbol{B}(k)$, and particles $(X_k^{\boldsymbol{i}})_{\boldsymbol{i}\in S_k}$. For all $\boldsymbol{i}\in S_k\backslash\boldsymbol{B}(k)$, we set $N_{k+1}^{\boldsymbol{i}}=\varphi_{k+1}(G_k(X_k^{\boldsymbol{i}}),V_{\boldsymbol{i}})$, and for $j\in[N_k^{\boldsymbol{i}}]$, we set $X_{k+1}^{(\boldsymbol{i},j)}=m_{k+1}(X_k^{\boldsymbol{i}},U_{(\boldsymbol{i},j)})$. We set $N_{k+1}^{\boldsymbol{B}(k)}=\widehat{\varphi}_{k+1}(G_k(X_k^{\boldsymbol{B}(k)}),V_{\boldsymbol{i}})$, $X_{k+1}^{\boldsymbol{B}(k+1)}=x_{k+1}$, and for $j\in\{2,\dots,N_{k+1}^{\boldsymbol{B}(k)}\}$, we set $X_{k+1}^{(\boldsymbol{B}(k),j)}=m_{k+1}(X_k^{\boldsymbol{B}(k)},U_{(\boldsymbol{B}(k),j)})$. We set $S_{k+1}=\cup_{\boldsymbol{i}\in S_k}\cup_{j\in[N_{k+1}^{\boldsymbol{i}}]}(\boldsymbol{i},j)$. We set $N_{k+1}=\#S_{k+1}$.

We now introduce an assumption about the existence of a dominating potential.

Hypothesis 1. *At each time-step $k\in[T-1]$, there exists a dominating potential $\widetilde{G}_k:\mathbb{N}^k\times E_k\times\mathcal{O}\to\mathbb{R}$ such that, for all $x\in E_k$, $\boldsymbol{i}\in\mathbb{N}^k$, $\Theta\in\mathcal{O}$, $G_k(x)\le\widetilde{G}_k(\boldsymbol{i},x,\Theta)\le\|G_k\|_\infty$.*

For further use, using the same random variables as the ones we used above, we can build a bigger conditioner forest, again recursively.

- We take the same $\boldsymbol{B}$ as above. For $i\in[n_1]$, we set $\widetilde{X}_1^i=X_1^i$. We set $\widetilde{S}_1=S_1$.
- Suppose we have $\widetilde{S}_k\subset(\mathbb{N}^*)^k$ of cardinality $\widetilde{N}_k$, containing $\boldsymbol{B}(k)$, and particles $(\widetilde{X}_k^{\boldsymbol{i}})_{\boldsymbol{i}\in S_k}$. For all $\boldsymbol{i}\in\widetilde{S}_k\backslash\boldsymbol{B}(k)$, we set

$$\widetilde{N}_{k+1}^{\boldsymbol{i}}=\begin{cases}\varphi_{k+1}(\widetilde{G}_k(\boldsymbol{i},\widetilde{X}_k^{\boldsymbol{i}}),V_{\boldsymbol{i}},\Theta) & \text{if } \boldsymbol{i} \text{ is a descendant of } \boldsymbol{B}(1),\,\boldsymbol{B}(2),\,\dots\text{or } \boldsymbol{B}(k),\\ \varphi_{k+1}(G_k(\widetilde{X}_k^{\boldsymbol{i}}),V_{\boldsymbol{i}}) & \text{otherwise,}\end{cases}\tag{4.12}$$

and for $j\in[\widetilde{N}_k^{\boldsymbol{i}}]$, we set $\widetilde{X}_{k+1}^{(\boldsymbol{i},j)}=m_{k+1}(\widetilde{X}_k^{\boldsymbol{i}},U_{(\boldsymbol{i},j)})$. We set

$$\widetilde{N}_{k+1}^{\boldsymbol{B}(k)}=\widehat{\varphi}_{k+1}(\widetilde{G}_k(\boldsymbol{B}(k),\widetilde{X}_k^{\boldsymbol{B}(k)},\Theta),V_{\boldsymbol{B}(k)})\,,$$

$\widetilde{X}_{k+1}^{\boldsymbol{B}(k+1)}=x_{k+1}$, and for $j\in\{2,\dots,\widetilde{N}_{k+1}^{\boldsymbol{B}(k)}\}$, we set $\widetilde{X}_{k+1}^{(\boldsymbol{B}(k),j)}=m_{k+1}(\widetilde{X}_k^{\boldsymbol{B}(k)},U_{(\boldsymbol{B}(k),j)})$. We set $\widetilde{S}_{k+1}=\cup_{\boldsymbol{i}\in\widetilde{S}_k}\cup_{j\in[N_{k+1}^{\boldsymbol{i}}]}(\boldsymbol{i},j)$. We set $\widetilde{N}_{k+1}=\#\widetilde{S}_{k+1}$. The bigger potentials $\widetilde{G}_k$ are used only on $\boldsymbol{B}(1)$, $\boldsymbol{B}(2)$, ... and their descendance. We use the potentials G_k on the other particles.

One can show recursively on k that for all $k\in[T]$, $S_k\subset\widetilde{S}_k$, for all $\boldsymbol{i}\in S_k$, $X_k^{\boldsymbol{i}}=\widetilde{X}_k^{\boldsymbol{i}}$, $N_{k+1}^{\boldsymbol{i}}\le\widetilde{N}_{k+1}^{\boldsymbol{i}}$ (almost surely in ω). We then build a proposal forest in a similar way, recursively on k.

- We set $\overline{S}_1=[n_1]$. For $\boldsymbol{i}\in S_1$, We set $\overline{X}_1^{\boldsymbol{i}}=m_1(U'_{\boldsymbol{i}})$.
- Suppose we have $\overline{S}_k\subset(\mathbb{N}^*)^k$ of cardinality $\overline{N}_k$ and particles $(\overline{X}_k^{\boldsymbol{i}})_{\boldsymbol{i}\in S_k}$. For all $\boldsymbol{i}\in\overline{S}_k$, we set $\overline{N}_{k+1}^{\boldsymbol{i}}=\varphi_{k+1}(G_k(\overline{X}_k^{\boldsymbol{i}}),V'_{\boldsymbol{i}})$, and for $j\in[\overline{N}_k^{\boldsymbol{i}}]$, we set $\overline{X}_{k+1}^{(\boldsymbol{i},j)}=m_{k+1}(\overline{X}_k^{\boldsymbol{i}},U'_{(\boldsymbol{i},j)})$. We set $\overline{S}_{k+1}=\cup_{\boldsymbol{i}\in S_k}\cup_{j\in[\overline{N}_{k+1}^{\boldsymbol{i}}]}(\boldsymbol{i},j)$. We set $\overline{N}_{k+1}=\#\overline{S}_{k+1}$.

Then we order $\overline{S}_T$ alphanumerically: $\overline{S}_T=\{\boldsymbol{i}_1,\dots,\boldsymbol{i}_{\overline{N}_T}\}$. We set

$$\boldsymbol{B}_T^\star=\boldsymbol{i}_r \text{ if } W_1\in\left[\frac{r-1}{\overline{N}_T},\frac{r}{\overline{N}_T}\right)\,,\ r\in[\overline{N}_T]\,.$$

We do not need to define $\boldsymbol{B}_T^*$ in the case $\overline{N}_T=0$. The accept/reject step then goes in the following way:

$$\text{if } W_2\le\min\left(1,\frac{\overline{N}_T}{N_T}\right),\text{ move to }(\overline{X}_1^{\boldsymbol{B}^\star(1)},\dots,\overline{X}_T^{\boldsymbol{B}^\star(T)})\,,$$

otherwise, stay in $(x_1,\dots,x_T)$. We will sometimes insist on the dependence of the variables on Θ, $x_1,\dots,x_T$ by writing $\widetilde{N}_T=\widetilde{N}_T(\Theta,(x_1,\dots,x_T))$, $N_T=N_T(\Theta)$, $W_2=W_2(\Theta)$, ...

Remark 4.3. For each of the branching processes in this Section, for all $k,q\in[T]$, $\boldsymbol{i}\in(\mathbb{N}^*)^k$, $\boldsymbol{j}$ in $(\mathbb{N}^*)^q$, if $\boldsymbol{i}$ is not an ancestor of $\boldsymbol{j}$ and $\boldsymbol{j}$ is not an ancestor of $\boldsymbol{i}$, then, conditionally on $(X_k^{\boldsymbol{i}},X_q^{\boldsymbol{j}})$ (resp.

$(\widetilde{X}_k^{\boldsymbol{i}}, \widetilde{X}_q^{\boldsymbol{j}})$, $(\overline{X}_k^{\boldsymbol{i}}, \overline{X}_q^{\boldsymbol{j}}))$, the descendants of $X_j^{\boldsymbol{i}}$ (resp. $\widetilde{X}_k^{\boldsymbol{i}}$, $\overline{X}_k^{\boldsymbol{i}}$) are independent of the descendants of $X_q^{\boldsymbol{j}}$ (resp. $\widetilde{X}_q^{\boldsymbol{j}}$, $\overline{X}_q^{\boldsymbol{j}}$). In the same way, when we sample a conditional forest (resp. a bigger conditional forest) conditionally to $(x_1, \dots, x_T)$, for all $k \in [T-1]$, the descendants of $X_k^{\boldsymbol{B}(k)}$ (resp. $\widetilde{X}_k^{\boldsymbol{B}(k)}$) with indexes $\boldsymbol{i} \notin \{\boldsymbol{B}(k+1), \dots, \boldsymbol{B}(T)\}$ depend only on Θ and x_k.

4.3. **Backward coupling.** Suppose we are given i.i.d. random variables $(\Theta_0, \Theta_1, \Theta_2, \dots)$ having the same law as Θ (all of them are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Any of these random variables is sufficient to perform a simulation of the Markov transition Q. The following result is a consequence of Theorem 3.1 of [FT98] (the original result can be found in [PW96]).

Theorem 4.4. *Suppose we have a function*

$$F \,:\, (\theta, x) \in \mathcal{O} \times (E_1 \times \cdots \times E_T) \mapsto F_\theta(x) \in E_1 \times \cdots \times E_T\,,$$

such that, for all x in $E_1 \times \cdots \times E_T$,

$$\theta \mapsto F_\theta(x)$$

is measurable and $F_\Theta(x)$ is of law $Q(x,.)$. If τ is a stopping time with respect to the filtration $(\sigma(\Theta_0, \dots, \Theta_n))_{n \geq 0}$ such that for all $(x_1, \dots, x_T)$, $(x_1', \dots, x_T')$ in $E_1 \times \cdots \times E_T$,

$$F_{\Theta_0} \circ F_{\Theta_1} \circ \cdots \circ F_{\Theta_\tau}(x_1, \dots, x_T) = F_{\Theta_0} \circ F_{\Theta_1} \circ \cdots \circ F_{\Theta_\tau}(x_1', \dots, x_T')\,,$$

then, for any $(x_1, \dots, x_T)$ in $E_1 \times \cdots \times E_T$,

$$F_{\Theta_0} \circ F_{\Theta_1} \circ \cdots \circ F_{\Theta_\tau}(x_1, \dots, x_T) \text{ is of law } \pi\,.$$

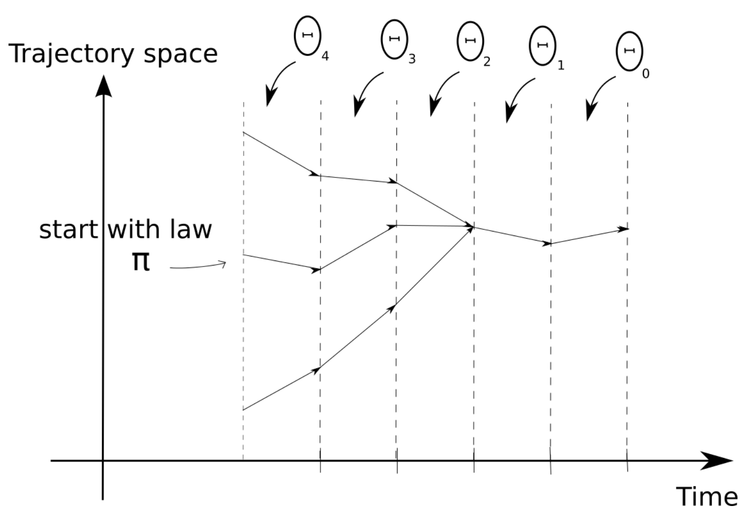

FIGURE 4.1. Backward coupling

In Figure 4.1, we draw an illustration of the above Theorem. The vertical axis represents the trajectory space $E_1 \times \cdots \times E_T$. For a realisation of $(\Theta_0, \Theta_1, \dots)$, τ is equal to 4. For various points $\xi(-1)$ in $E_1 \times \cdots \times E_T$, we draw the trajectories $\xi(0) = F_{\Theta_4}(\xi(-1))$, ..., $\xi(4) = F_{\Theta_0}(\xi(3))$. By definition of τ, the endpoint $\xi(4)$ does not depend on $\xi(-1)$.

We suppose we have dominating potentials $\widetilde{G}_k$, $k \in [T-1]$ as in Subsection 4.2 above. We write for all n, $\Theta_n = (U_{n,\boldsymbol{i}}, U'_{n,\boldsymbol{i}}, V'_{n,\boldsymbol{i}}, V_{n,\boldsymbol{i}}, W_{n,1}, W_{n,2})_{k \in \mathbb{N}^*, \boldsymbol{i} \in (\mathbb{N}^*)^k}$. In our simulations, we use the following stopping time

$$\begin{aligned}(4.13)\quad \tau &= \min\left\{ n \,:\, W_{n,2} \leq \min\left(1, \frac{\overline{N}_T(\Theta_n)}{\widetilde{N}_T(\Theta_n, (x_1, \dots, x_T))}\right), \forall (x_1, \dots, x_T) \in E_1 \times \cdots \times E_T \right\} \\ &= \min\left\{ n \,:\, W_{n,2} \leq \min\left(1, \frac{\overline{N}_T(\Theta_n)}{\sup_{(x_1,\dots,x_T) \in E_1 \times \cdots \times E_T} \widetilde{N}_T(\Theta_n, (x_1, \dots, x_T))}\right) \right\}\end{aligned}$$

Algorithm 1 Exact simulation

for n in $\mathbb{N}$ repeat until $n = \tau$
    draw Θ_n, and store it
    test whether $n = \tau$ or not
pick any trajectory $(x_1, \dots, x_T) \in E_1 \times \cdots \times E_T$
set $\xi(-1) = (x_1, \dots, x_T)$
for $n = 0$ to τ repeat
    $\xi(n) = F_{\Theta_{\tau-n}}(\xi(n-1))$
return $\xi(\tau)$

This stopping time satisfies the assumptions of the above Theorem. Algorithm 1 is thus an exact simulation of the law π. At this point, this algorithm is merely theoretical. The following two remarks will make it implementable, at least in some cases.

- We need to be able to compute $\sup\{\widetilde{N}_T(\Theta(\omega), (x_1, \dots, x_T)), (x_1, \dots, x_T) \in E_1 \times \cdots \times E_T\}$ for a fixed ω. The easiest case is where E_1, E_2, $\dots$, E_T are finite. We will see below how to reduce the problem to this case in cases where E_1, E_2 $\dots$ are not finite. If E_1 is finite, we can look for $x_1 \in E_1$ maximizing the descendants of $\widetilde{X}_1^1$ at time T (using $\Theta(\omega)$ to make the simulation), and so on. As we said in Remark 4.3, once $\widetilde{X}_1^1, \dots, \widetilde{X}_T^{(1,\dots,1)}$ are fixed, their descendants are independent, this is what makes the use of branching processes interesting.
- In Algorithm 1, we first sample $\Theta_0(\omega), \Theta_1(\omega), \dots$ until τ. And then we need the same realisations of the variables $(\Theta_0(\omega),\ \Theta_1(\omega),\ \dots)$ to compute $\xi(0),\ \dots,\ \xi(\tau)$. The object $\Theta_0(\omega)$ is an infinite collection of numbers so it is impossible to store. We set $\mathcal{E}(\omega)$ to be the subset of indexes $\boldsymbol{i} \in \cup_{n\in[T]}(\mathbb{N}^*)^n$ such that $U_{n,\boldsymbol{i}}(\omega)$ or $V_{n,\boldsymbol{i}}(\omega)$ is used when computing $\sup_{(x_1,\dots,x_T)\in E_1\times\cdots\times E_T} \widetilde{N}_T(\Theta_n(\omega), (x_1, \dots, x_T))$. We remark that, for all n, we do not need to store the whole $\Theta_n(\omega)$; having stored the number of descendants at time T of $\widetilde{X}_1^2(\omega), \dots, \widetilde{X}_1^{n_1}(\omega)$ (these are the starting points in the building of the "bigger conditional forest above"), $\overline{N}_T(\Theta_n(\omega))$, $(\overline{X}_1^{\boldsymbol{B}^\star(1)}, \dots, \overline{X}_T^{\boldsymbol{B}^\star(T)})(\omega)$ (this is the coloured trajectory in the proposal above) and

$$\{(U_{\boldsymbol{i}}(\omega), V_{\boldsymbol{i}}(\omega), W_2(\omega))_{\boldsymbol{i}\in(\mathbb{N}^*)^n, n\in[T]} : \boldsymbol{i}(1) = 1, \boldsymbol{i} \in \mathcal{E}(\omega)\} \tag{4.14}$$

  is enough to compute $F_{\Theta_n(\omega)}(\xi)$ for any ξ in $E_1 \times \cdots \times E_T$. The collection of numbers in (4.14) contains the number which might be needed when we compute the descendance of $X_1^1(\omega)$, $X_2^1(\omega)$, $\dots$, $X_{T-1}^1(\omega)$ in what is called above the "bigger conditional forest", and we do not need any other numbers. Another point is that we can code the simulation in such a way that we sample the variables in (4.14) in the alphanumerical order of their indexes $\boldsymbol{i}$ at each time step. So, instead of storing these variables, we can store random seeds. For example, instead of storing $U^{(1,1,1)}(\omega)$, $U^{(1,1,2)}(\omega)$, $U^{(1,1,3)}(\omega)$, $\dots$, we can store a single random seed[1].

We are now able to explain the purpose of Subsection 4.1. It is clear that when simulating a branching process, whether it is a conditional forest or a proposal forest, we do not want the number of particles to grow up. Such a growth would be exponential in T, which would be very bad for the complexity of our algorithm. On the other hand, if our branching processes become extinct before time T, then $\overline{N}_T(\Theta_n) = 0$, leading to $\tau \neq n$, and thus the first loop of Algorithm 1 could go on for a very long time. Again, this would be very bad for the complexity of our algorithm.

[1]We recall here that when the user asks for random variables $U^{(1,1,1)}(\omega)$, $U^{(1,1,2)}(\omega)$, $U^{(1,1,3)}(\omega)$, $\dots$, a computer will return numbers picked in a non-random list. So instead of storing theses random variables, we can store only the starting point in the list (the so-called "random seed").

5. Examples

5.1. **Self-avoiding random walks in $\mathbb{Z}^2$.** There exists an algorithm for simulating self-avoiding random walks in $\mathbb{Z}^2$ with approximatively the desired law (this law is described below). It is a MCMC (Monte-Carlo Markov Chain) algorithm. See [Ken02] and the references therein.

5.1.1. *Description of the model.* We take $E_1 = \mathbb{Z}^2$, $E_2 = \mathbb{Z}^2 \times \mathbb{Z}^2$, $\dots$, $E_T = (\mathbb{Z}^2)^T$ and for all $n \in [T]$, $(z_1, \dots, z_n) \in (\mathbb{Z}^2)^n$,

$$G_n(z_1, \dots, z_n) = \begin{cases} 1 & \text{if } z_i \neq z_j \text{ for all } i, j \in [n],\ i \neq j, \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

We take $M_1 = \delta_{(0,0)}$ (the Dirac mass at the origin of $\mathbb{Z}^2$). For all $n \in [T-1]$, $(z_1, \dots, z_n) \in (\mathbb{Z}^2)^n$, we take

$$\begin{aligned} &M_{n+1}((z_1, \dots, z_n), (z_1, \dots, z_n, z_n + (0,1)) = M_{n+1}((z_1, \dots, z_n), (z_1, \dots, z_n, z_n + (0,-1)) \\ &= M_{n+1}((z_1, \dots, z_n), (z_1, \dots, z_n, z_n + (1,0)) = M_{n+1}((z_1, \dots, z_n), (z_1, \dots, z_n, z_n + (-1,0)) = \frac{1}{4}\,. \end{aligned} \tag{5.2}$$

Then the marginal of π on E_{T-1} is the uniform law on the set of paths $(z_1, z_2, \dots, z_{T-1}) \in \mathbb{Z}^2$ such that $z_1 = (0,0)$, $|z_i - z_{i+1}| = 1$ for all i ($|\dots|$ being the Euclidean norm), for $i, j \in [T-1]$ with $i \neq j$, $z_i \neq z_j$ (the path does not intersect with itself, one also says that it is self-avoiding).

5.1.2. *Stopping time.* We set $\boldsymbol{B} = (1, 1, \dots, 1) \in \mathbb{N}^T$. For $k \in [T]$, $\boldsymbol{i} \in \mathbb{N}^k$, $q \in [k]$ such that $\boldsymbol{i}(q) = \boldsymbol{B}(q)$ and $\boldsymbol{i}(q+1) \neq \boldsymbol{B}(q+1)$, we set, for all $x = (z_1, \dots, z_k) \in E_k$, $\Theta \in \mathcal{O}$, $\widetilde{G}_k(\boldsymbol{i}, x, \Theta) = G_{k-q}(z_q, z_{q+1}, \dots, z_k)$ in other words

$$\widetilde{G}_k(\boldsymbol{i}, x) = \begin{cases} 1 & \text{if } (z_q, \dots, z_k) \text{ is self-avoiding,} \\ 0 & \text{otherwise} \end{cases}$$

(as $\widetilde{G}_k$ does not depend on Θ in this example, we replace $\widetilde{G}_k(\boldsymbol{i}, x, \Theta)$ by $\widetilde{G}_k(\boldsymbol{i}, x)$). We do not need to define $\widetilde{G}$ in the remaining cases. As we said in Subsection 4.3, we sample variables Θ_0, Θ_1, $\dots$ and we look for the stopping time τ defined in (4.13). For fixed n, Θ_n, and $k \in [T-1]$, $x_k \in E_k$, if we sample a bigger conditional forest with $\widetilde{X}_k^{\boldsymbol{B}(k)} = x_k$, we introduce the following notation:

$$\widetilde{N}_T(\Theta_n, x_k) = \#\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \boldsymbol{i} \neq \boldsymbol{B}(T), \boldsymbol{i}(k+1) \neq 1, \boldsymbol{B}(k) \prec \boldsymbol{i}\}\,.$$

We do not need the values $\widetilde{X}_q^{\boldsymbol{B}(q)}$ for $q \neq k$ to compute the above quantity. Due to the form of the potentials $\widetilde{G}_.$, the set $\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \boldsymbol{i} \neq \boldsymbol{B}(T), \boldsymbol{B}(k) \prec \boldsymbol{i}\}$ depends only on Θ_n, and

$$\text{for all } x_k \in (\mathbb{Z}^2)^k,\ \widetilde{N}_T(\Theta_n, x_k) = \widetilde{N}_T(\Theta_n, ((0,1), \dots, (0,k)))\,.$$

We set $\widetilde{\widetilde{N}}_T(\Theta_n, k) = \widetilde{N}_T(\Theta_n, ((0,1), \dots, (0,k)))$. Now we have, for all $(x_1, \dots, x_T) \in E_1 \times \dots \times E_T$,

$$\widetilde{N}_T(\Theta_n, (x_1, \dots, x_T)) = \widetilde{\widetilde{N}}_T(\Theta_n) := \#\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \nexists k : \boldsymbol{B}(k) \prec \boldsymbol{i}\} + 1 + \sum_{k=1}^{T-1} \widetilde{\widetilde{N}}_T(\Theta_n, k)\,. \tag{5.3}$$

This equation means that the supremum in (4.13) is easy to find. And so, by Remark 4.3,

$$\tau = \min\left\{ n \geq 0 : W_{n,2} \leq \min\left(1, \frac{\overline{N}_T(\Theta_n)}{\widetilde{\widetilde{N}}_T(\Theta_n)}\right)\right\}\,.$$

5.1.3. *Complexity of the algorithm: a case study.* We take here the following law for the simulation of the number of children

$$\text{for all } k,\ f_{k+1}(0,0) = 0\,,\ f_{k+1}(1,1) = p_{k+1}\,,\ f_{k+1}(1,2) = 1 - p_{k+1}\,,$$

for some sequences $(p_k)_{2 \leq k \leq T}$ taking values in $(0,1)$. We now look at a branching process in E_1, E_2, $\dots$,E_T based on the potential defined in (5.1), the transitions defined in (5.2) and the above reproduction law. A sensible way of choosing the constants (p_k)'s is to choose them such that a branching process starting with n_1 particles will have a random number N_k of descendants of

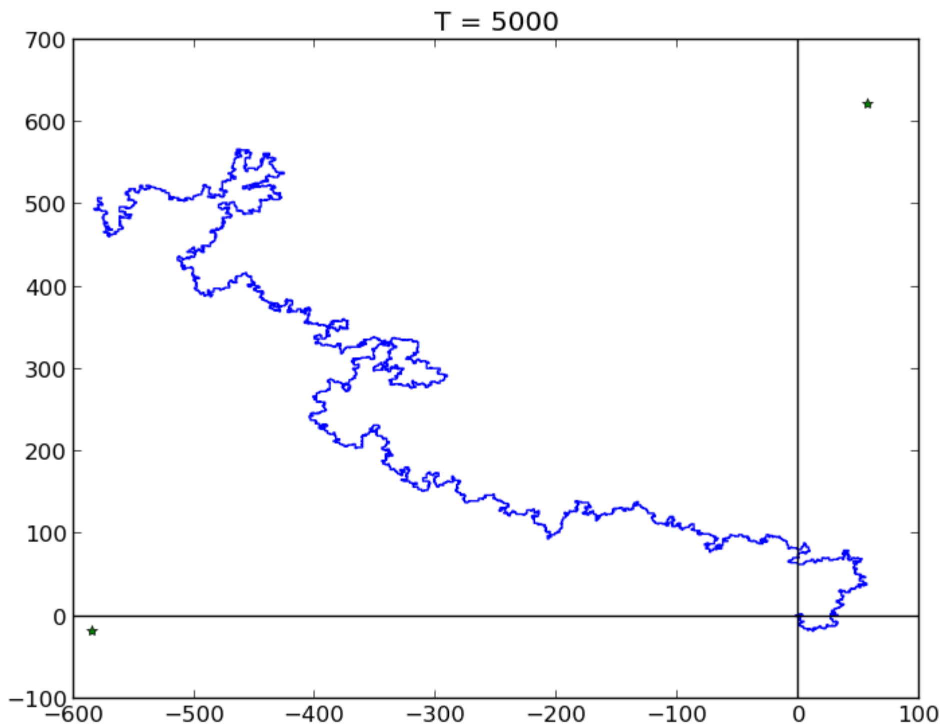

FIGURE 5.1. Self-avoiding random walk in $\mathbb{Z}^2$.

T	100	200	300	350	400
$C(T)$	14.90	28.45	19.89	23.28	30.48

TABLE 1. $C(T)$ for the self-avoiding random walk

the same order of magnitude as n_1 (this requires some pre-processing). These numbers N_k are random but the law of large numbers makes them not too fluctuant. It turns out that a good tuning is to have the (p_k)'s almost constant. By doing so, we are able to draw a trajectory with law π by a matter of minutes if $T \leq 1000$ and by a matter of one hour if $T \leq 5000$ (see Figure 5.1). Here, we ran a program in `C`. We used parallelization to make it faster (with the `OpenMP` library). The program uses around five cores simultaneously. Laptop computers are multi core nowadays, so the limiting factor is not the number of cores but the management of the memory. Indeed, the genealogies we build in our algorithm can take a lot of space, if the code is not written properly. An basic calculation shows that n_1 should be chosen as

$$n_1(T) = \sup \left(\frac{16\sigma_1(T)^2}{\mu_1(T)}, \frac{\mu_2(T)}{\mu_1(T)} \right) ,$$

where $\mu_1(T)$, $\mu_2(T)$ are expectations, $\sigma_1^2(T)$ is a variance, these terms being defined in Section 6.2 (see Equations (6.4), (6.9)). The quantity $\mu_1(T)$ is constant (equal to 1) and $\sigma_1^2(T)$ is expected to be polynomial in T. So we write $n_1(T) = C_0 + C(T) \times T$, with C_0 a constant and $C(T)$ depending on T. We estimate $C(T)$ by Monte-Carlo for $T \in \{100, 200, 300, 350, 400\}$ (see the Appendix for details, we use 1000 samples for each expectation and variance we have to estimate). We can then compare T and $C(T)$ (see Table 1). A simple least square regression in log-log scale gives a slope of 0.27. So it seems sensible to take n_1 proportional to T or $T^{3/2}$.

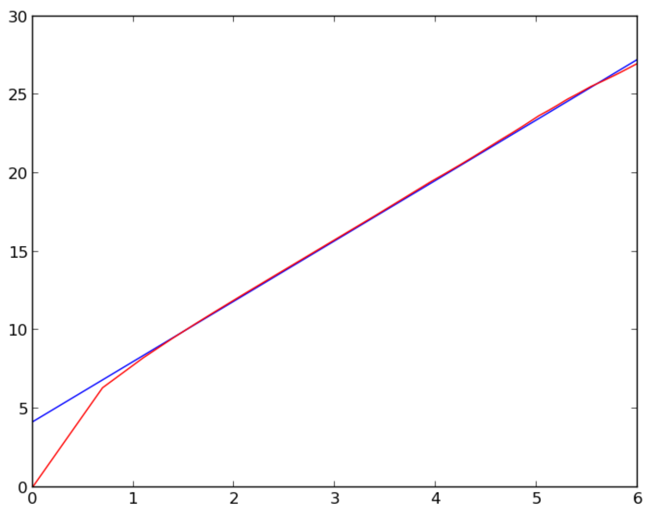

FIGURE 5.2. Log-log graph of complexity versus time in the case of the self-avoiding random walk.

We then want to estimate the average number of particles at each generation when we run a simulation of a bigger conditional forest, assuming, we take $n_1 = T^{3/2}$. For a fixed T, the average complexity of drawing one sample of a bigger conditional forest is the sum on all generations of the average number of particles at each generation times the number of the generation (this is the cost of computing the potential). With this choice of dependency between n_1 and T, this complexity is the average complexity of our algorithm (see Remark 6.4).

Using a Monte-Carlo method with 1000 samples for each expectation we have to estimate, we are able to draw in Figure 5.2 the log of the expected complexity against $\log(T)$ ($T \in [400]$) with $n_1 = T^{3/2}$. We draw a linear regression on the same graph. The estimated slope is 3.85. So the complexity seems to be polynomial in T.

5.2. **Filter in $\mathbb{R}^3$.**

5.2.1. *Description of the model.* We are given the following signal/state ($(X_n)_{n\geq 1}$) and observations ($(Y_n)_{n\geq 1}$) in $\mathbb{R}^3$:

$$\begin{cases} X_{n+1} &= AX_n + V_{n+1}\,, \quad \forall n \geq 1\,, \\ Y_n &= X_n + W_n\,, \quad \forall n \geq 1\,, \end{cases}$$

with X_1 following a law M_1 and $(V_n)_{n\geq 2}$ independent of $(W_n)_{n\geq 1}$, the V_n's are i.i.d. with a law of density f and the W_n's are i.i.d. with a law of density g (with respect to the Lebesgue measure). The coefficient A is a 3×3 real matrice. We suppose we have functions F and G such that, for all $U \in [0,1]$, $F(U)$ is a random variable in $\mathbb{R}^3$ of law of density f, $G(U)$ is a random variable in $\mathbb{R}^3$ of law of density g.

We are interested in $\mathcal{L}(X_1,\dots,X_T|Y_1,\dots,Y_{T-1})$ for some $T \in \mathbb{N}^*$. From now on, we will suppose that the sequence Y_1, Y_2, ... is fixed. In particular, all expectations will be conditional to $Y_{1:T-1}$. We set, for all $k \in \mathbb{N}^*$, $G_k(x) = g(Y_k - x)$. We denote by $M_2 = M_3 = \cdots = M$ the transition kernel of the Markov chain $(X_n)_{n\geq 1}$. We set $E_1 = E_2 = \cdots = \mathbb{R}^3$. Then $\mathcal{L}(X_1,\dots,X_T|Y_1,\dots,Y_{T-1})$ coincides with π defined in (1.1). We make the following hypotheses.

Hypothesis 2. *The matrix A is invertible. For all $x, y \in \mathbb{R}^3$, $|Ax - Ay| \leq \alpha|x-y|$ ($|\dots|$ is the Euclidean norm) with $\alpha \in [0,1)$.*

Hypothesis 3. *We have $g(x) \underset{|x|\to+\infty}{\longrightarrow} 0$.*

5.2.2. *Computing the stopping time.* We take m_k introduced in (4.11) to be, for all $x \in \mathbb{R}^3$, $U \in [0,1]$,

$$m_k(x,U) = Ax + F(U)\,.$$

We fix $\delta > 0$. For $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, we set

$$L_\delta(x) = \left[\delta\left\lfloor\frac{x_1}{\delta}\right\rfloor, \delta\left\lfloor\frac{x_1}{\delta}\right\rfloor + \delta\right) \times \left[\delta\left\lfloor\frac{x_2}{\delta}\right\rfloor, \delta\left\lfloor\frac{x_2}{\delta}\right\rfloor + \delta\right) \times \left[\delta\left\lfloor\frac{x_3}{\delta}\right\rfloor, \delta\left\lfloor\frac{x_3}{\delta}\right\rfloor + \delta\right)\,.$$

We set $\boldsymbol{B} = (1,1,\dots,1) \in \mathbb{N}^T$. We suppose we are given a random variable Θ as in Subsection 4.2. We consider $k \in [T]$, $x \in \mathbb{R}^3$, $\boldsymbol{i} \in \mathbb{N}^k$ such that there exists $q \in [k]$ satisfying $\boldsymbol{i}(q) = \boldsymbol{B}(q)$, $\boldsymbol{i}(q+1) \neq \boldsymbol{B}(q+1)$. There exists one and only one sequence $(x_q, x_{q+1}, \dots, x_k)$ such that

$$\begin{aligned} x_{q+1} &= m_{q+1}(x_q, U_{\boldsymbol{i}(q+1)})\,, \\ x_{q+2} &= m_{q+2}(x_{q+1}, U_{\boldsymbol{i}(q+2)})\,, \\ &\dots \\ x = x_k &= m_k(x_{k-1}, U_{\boldsymbol{i}})\,. \end{aligned}$$

For $y \in \mathbb{R}^3$, we introduce:

$$\begin{aligned} m_{q,q+1}(y) &= m_{q+1}(y, U_{\boldsymbol{i}(q+1)})\,, \\ m_{q,q+2}(y) &= m_{q+2}(m_{q,q+1}(y), U_{\boldsymbol{i}(q+2)})\,, \\ &\dots \\ m_{q,k}(y) &= m_k(m_{q,k-1}(y), U_{\boldsymbol{i}})\,. \end{aligned}$$

We set

$$\widetilde{G}_k(\boldsymbol{i}, x, \Theta) = \sup_{y \in L_\delta(x_q)} G_k(m_{q,k}(y))\,. \tag{5.4}$$

This implies that $G_k(x) \leq \widetilde{G}_k(\boldsymbol{i}, x, \Theta)$. The idea here is to bound the potential $G_k(x)$ by its supremum on a subset of $\mathbb{R}^3$ containing x. Due to Hypothesis 2, the diameter of $\{m_{q,k}(y) : y \in L_\delta(x_q)\}$ in (5.4) is bounded by $(\delta\sqrt{3})^{k-q}$. Under the additional assumption that g is continuous, it will make that $G_k(x)$ is not too far from $\widetilde{G}_k(\boldsymbol{i}, x, \Theta)$ in the above bound. And so, the number of descendants of $\widetilde{X}_k^{\boldsymbol{B}(k)}$ should not explode when $T-k$ becomes big. These are only heuristics and we will study the complexity of the algorithm based on these $\widetilde{G}_.$ below. For the sake of completeness, we define, for k in $[T]$ for x in E_k, $\boldsymbol{i}$ in $\mathbb{N}^k$ such that there exists no q in $[k]$ such that $\boldsymbol{i}(q) = \boldsymbol{B}(q)$,

$$\widetilde{G}_k(\boldsymbol{i}, x, \Theta) = G_k(x)\,.$$

As we said in Subsection 4.3 and in the previous example, we sample variables Θ_0, Θ_1, … and we look for the stopping time τ defined in (4.13). For fixed n, Θ_n and $k \in [T-1]$, $x_k \in \mathbb{R}^3$, we sample a bigger conditional forest with $\widetilde{X}_k^{\boldsymbol{B}(k)} = x_k$ (this bigger conditional forest is based on the dominating potentials $\widetilde{G}_k$, as in Section 4.3). We introduce the following notation

$$\widetilde{N}_T(\Theta_n, k, x_k) = \#\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \boldsymbol{i} \neq \boldsymbol{B}(T), \boldsymbol{B}(k) \prec \boldsymbol{i}\}\,.$$

We do not need the values $\widetilde{X}_q^{\boldsymbol{B}(q)}$, $q \neq k$ to compute the above quantity.

Let us look at Figure 5.3. We have a realisation of a variable of a branching process using potentials $(\widetilde{G}_k)_{k \in [T-1]}$ (built with some variable Θ). The coloured trajectory is coloured in red. We have here

$$\widetilde{N}_T(\Theta, 1, \widetilde{X}_T^{\boldsymbol{B}(1)}) = 0\,,\ \widetilde{N}_T(\Theta, 2, \widetilde{X}_T^{\boldsymbol{B}(2)}) = 4\,,\ \widetilde{N}_T(\Theta, 3, \widetilde{X}_T^{\boldsymbol{B}(3)}) = 0\,,\ \widetilde{N}_T(\Theta, 4, \widetilde{X}_T^{\boldsymbol{B}(4)}) = 1\,.$$

We define

$$\widetilde{\widetilde{N}}_T(\Theta_n, k) = \sup_{x_k \in \mathbb{R}^3} \widetilde{N}_T(\Theta_n, k, x_k)\,. \tag{5.5}$$

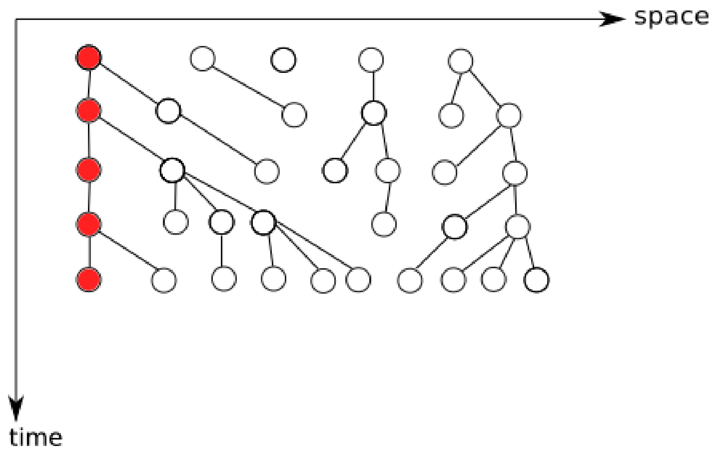

FIGURE 5.3. Bigger conditional forest

We define $\widetilde{\widetilde{N}}_T(\Theta_n)$ by

$$\widetilde{\widetilde{N}}_T(\Theta_n) := \sup_{(x_1,\dots,x_T)\in(\mathbb{R}^3)^T} \widetilde{N}_T(\Theta_n,(x_1,\dots,x_T))\,,$$

it satisfies

$$\widetilde{\widetilde{N}}_T(\Theta_n) = \#\{i\in\widetilde{S}_T(\Theta_n),\,\nexists k:\,\boldsymbol{B}(k)\prec\boldsymbol{i}\}+1+\sum_{k=1}^{T-1}\widetilde{\widetilde{N}}_T(\Theta_n,k)\,. \tag{5.6}$$

For fixed n, Θ_n and $k\in[T-1]$, suppose $x_k,x'_k\in\mathbb{R}^3$ are such that $L_\delta(x_k)=L_\delta(x'_k)$ then the descendants of $\widetilde{X}_k^{\boldsymbol{B}(k)}$ in the bigger conditional forest are the same whether $\widetilde{X}_k^{\boldsymbol{B}(k)}=x_k$ or $\widetilde{X}_k^{\boldsymbol{B}(k)}=x'_k$ (the bigger conditional forest is built with the dominating potential $\widetilde{G}_k$). Suppose we have $\Theta_n=(U_{n,\boldsymbol{i}},U'_{n,\boldsymbol{i}},V_{n,\boldsymbol{i}},V'_{n,\boldsymbol{i}},W_{n,1},W_{n,2})_{\boldsymbol{i}\in\mathbb{N}^n,n\geq1}$, as in Section 4.3. The number of children of $\widetilde{X}_k^{\boldsymbol{B}(k)}$ is $\widetilde{N}_{k+1}^{\boldsymbol{B}(k)}=\widehat{\varphi}_{k+1}(\widetilde{G}_k(\boldsymbol{B}(k),\widetilde{X}_k^{\boldsymbol{B}(k)},\Theta),V_{n,\boldsymbol{B}(k)})$, which is equal to zero if $|\widetilde{X}_k^{\boldsymbol{B}(k)}-Y_k|$ is big enough under Hypothesis 3. So, the number of operations needed to compute $\widetilde{N}_T(\Theta_n,k)$ is finite for all k, Θ_n; and $\mathbb{E}(\widetilde{N}_T(\Theta,k))$ is finite for all k. This implies that $\mathbb{E}(\widetilde{N}_T(\Theta,k))$ is finite for all k. So, once we are given Θ_0, Θ_1, … the stopping time τ can be computed in finite time.

5.2.3. *Complexity of the algorithm.* As explained in Section 6.2, the complexity of the algorithm depends on constants $\mu_{2,k}$ ($k\in[T-1]$) chosen such that $\mu_2,k\geq1/4$ and

$$\mathbb{P}(\widetilde{\widetilde{N}}_T(\Theta,k)\geq4\mu_{2,k})\leq\frac{1}{4T}\text{ for all }k\in[T-1]. \tag{5.7}$$

We suppose we are in the case of the Kalman filter. We suppose that the filters $(\pi_k)_{k\geq1}$ have all the same variance (which is the case if the Markov chain $(Y_k,\pi_k)_{k\in[T]}$ is stationnary). The case where the filters do not have the same variances can be treated with the same ideas, but with more complicated calculations.

Hypothesis 4. *The random variables (V_n) are of law $\mathcal{N}(0,\sigma^2 Id)$. The random variables (W_n) are of law $\mathcal{N}(0,s^2 Id)$. The matrix A satisfies $A=a\times Id$ ($a\in(-1,1)$). The law M_1 is chosen such that for all $k\geq1$, π_k is a Gaussian with covariance matrix $\sigma_\infty^2 Id$.*

We set $m_k^\pi=\mathbb{E}(X_k|Y_1,\dots,Y_{k-1})$ for all k. We have, by a basic computation,

$$Y_k-m_k^\pi\sim\mathcal{N}(0,(\sigma_\infty^2+s^2)\mathrm{Id})\,,\,\forall k\,, \tag{5.8}$$

with

$$\sigma_\infty^2=\frac{1}{2}\left(-s^2(1-a^2)+\sigma^2+\sqrt{(s^2(1-a^2)-\sigma^2)^2+4\sigma^2s^2}\right)\,,$$

and the $(Y_k - m_k^\pi)_{k\in[T]}$ are independent. We suppose that

$$\delta = \frac{C_\delta}{T^2} \wedge 1\,.$$

for some constant C_δ. We introduce $\epsilon \in [0, 2/(\sqrt{3}(e-1))]$ such that

$$\epsilon = \frac{C_\epsilon}{T}\,,$$

for some constant C_ϵ.

Hypothesis 5. *We suppose we are in the sub-case of Equation (4.2) and that the constants $(\alpha_k)_{2\leq k\leq T}$, $(q_k)_{2\leq k\leq T}$ are chosen such that*

$$\frac{\alpha_{k+1}(q_{k+1}+1)}{2}\pi_k(G_k) = 1, \textit{ for all } k \textit{ in } [T-1]\,,$$

with

$$\alpha_{k+1} = \frac{2}{\pi_k(G_k)}\left\lceil \frac{2\|G_k\|_\infty}{\pi_k(G_k)} \right\rceil^{-1}\,,$$
$$q_{k+1} = \left\lceil \frac{2\|G_k\|_\infty}{\pi_k(G_k)} \right\rceil - 1\,.$$

Under the above Hypothesis, we have

$$q_{k+1} \geq 1\,,\ \alpha_{k+1}\|G_k\|_\infty \in \left[\frac{2}{3}, 1\right]\,. \tag{5.9}$$

Proposition 5.1. *Under Hypothesis 4, 5, we can find constants $\mu_{2,k}$ such that, when $T\to+\infty$,*

$(\mu_{2,k})_{k\in[T-1]}$ satisfies Equation (5.7),

$$\frac{1}{T^{\gamma'}}\sum_{k=1}^{T-1}\mu_{2,k+1} \xrightarrow[T\to+\infty]{law} 0\,.$$

for some $\gamma' > 0$.

Before going into the proof of this Proposition, we need to prove an auxiliary Lemma. For all $k\in\{2,\dots,T-1\}$, we set Z_k to be a random variable of law π_k. We define, for all $k\in\{2,\dots,T-1\}$, Θ, x,

$$\boldsymbol{i}_0 = (1,1,\dots,1,2)\in\mathbb{N}^k\,,$$

$$\breve{N}_T(\Theta,k,x) = \mathbb{E}(\#\{\boldsymbol{i}\in\widetilde{S}_T(\Theta), \boldsymbol{i}_0 \prec \boldsymbol{i}\}|\widetilde{X}_k^{\boldsymbol{i}_0} = x)\,.$$

Lemma 5.2. *There exists n such that*

$$\sup_{k\in[T-1]}\mathbb{E}(\breve{N}_T(\Theta,k,Z_k)) = O(1)\,,\ a.s.$$

Proof. Let $k\in[T-1]$. We recall that, if $\widetilde{X}_k^{\boldsymbol{i}_0} = Z_k$, then its number of children is

$$\varphi(\widetilde{G}_k(\boldsymbol{B}(k), Z_k, \Theta), V_{\boldsymbol{i}_0})$$

and the positions of these children are $(m_{k+1}(Z_k, U_{(\boldsymbol{i}_0,i)}))_{1\leq i\leq\varphi(\widetilde{G}_k(\boldsymbol{i}_0,Z_k,\Theta),V_{\boldsymbol{i}_0})}$. We have, for all $k\in[T-2]$,

$$\begin{aligned}
&\mathbb{E}(\breve{N}_T(\Theta,k,Z_k)\mathbb{1}_{|Z_k-Y_k|>s^2\epsilon/\delta})\\
&\qquad = \mathbb{E}(\sum_{i=1}^{\varphi(\widetilde{G}_k(\boldsymbol{i}_0,Z_k,\Theta),V_{\boldsymbol{i}_0})}\breve{N}_T(\Theta,k+1,m_{k+1}(Z_k,U_{(\boldsymbol{i}_0,i)}))\mathbb{1}_{|Z_k-Y_k|>s^2\epsilon/\delta})\\
&\qquad\qquad = \mathbb{E}(\varphi(\widetilde{G}_k(\boldsymbol{i}_0,Z_k,\Theta),V_{\boldsymbol{i}_0})\breve{N}_T(\Theta,k+1,m_{k+1}(Z_k,U_{(\boldsymbol{i}_0,1)}))\mathbb{1}_{|Z_k-Y_k|>s^2\epsilon/\delta})
\end{aligned}$$

because $(\widetilde{X}_{k+1}^{(\boldsymbol{i}_0,i)})_{1\leq i\leq\varphi(\widetilde{G}_k(\boldsymbol{i}_0,Z_k,\Theta),V_{\boldsymbol{i}_0})}$ are i.i.d. conditionally on Z_k, $\varphi(\widetilde{G}_k(\boldsymbol{i}_0,Z_k,\Theta),V_{\boldsymbol{i}_0})$. So

$$\begin{aligned}
(5.10)\quad \mathbb{E}(\breve{N}_T(\Theta,k,Z_k)\mathbb{1}_{|Z_k-Y_k|>s^2\epsilon/\delta}) &\leq \mathbb{E}(\varphi(\widetilde{G}_k(\boldsymbol{i}_0,Z_k,\Theta),V_{\boldsymbol{i}_0})\left(\prod_{j=k+2}^{T} q_j\right)\mathbb{1}_{|Z_k-Y_k|>s^2\epsilon/\delta}) \\
&= \mathbb{E}\left(\alpha_{k+1}\frac{(q_{k+1}+1)}{2}\widetilde{G}_k(\boldsymbol{i}_0,Z_k,\Theta)\left(\prod_{j=k+2}^{T} q_j\right)\mathbb{1}_{|Z_k-Y_k|>s^2\epsilon/\delta}\right) \\
&\leq \frac{1}{\pi_k(G_k)}\sup_{z:|z-Y_k|\geq\frac{s^2\epsilon}{\delta}}\widetilde{G}_k(\boldsymbol{i}_0,z,\Theta)\left(\prod_{j=k+2}^{T} q_j\right) \\
&\leq \frac{1}{\pi_k(G_k)\,(2\pi s^2)^{3/2}}\exp\left(-\frac{1}{2s^2}\left(\frac{s^2\epsilon}{\delta}-\sqrt{3}\delta\right)_+^2\right)\left(\prod_{j=k+2}^{T} q_j\right).
\end{aligned}$$

For x such that $|x-Y_k|\leq s^2\epsilon/\delta$, we have, for all $y\in L_\delta(x)$ ($\langle .,.\rangle$ is the standard scalar product)

$$\begin{aligned}
-\frac{|y-Y_k|^2}{2s^2} &= -\frac{|y-x|^2}{2s^2}-\frac{|x-Y_k|^2}{2s^2}-\frac{\langle y-x,x-Y_k\rangle}{2s^2} \\
&\leq -\frac{|x-Y_k|^2}{2s^2}+\frac{\sqrt{3}\delta(s^2\epsilon/\delta)}{2s^2} \\
&= -\frac{|x-Y_k|^2}{2s^2}+\frac{\sqrt{3}\epsilon}{2}.
\end{aligned}$$

And so (as $\epsilon\sqrt{3}/2\leq 1$)

$$G_k(y)\leq G_k(x)\exp(\sqrt{3}\epsilon/2)\leq G_k(x)\left(1+\frac{\sqrt{3}(e^1-e^0)\epsilon}{2}\right),$$

$$\widetilde{G}_k(x)\leq G_k(x)(1+\epsilon').$$

We set $\epsilon'=\sqrt{3}(e^1-e^0)\epsilon/2$ (we have $\epsilon'\in[0,1]$).

For Z_k such that $|Z_k-Y_k|\leq s^2\epsilon/\delta$, we introduce new variables.

- If $\alpha_{k+1}G_k(Z_k)\times(1+\epsilon')\leq 1$, we set

$$Z^i_{k+1}=m_{k+1}(Z_k,U_{(\boldsymbol{B}(k),i)}),\ \text{for } i\in\{1,\dots,\varphi_{k+1}(\alpha_{k+1}G_k(Z_k)(1+\epsilon'),V_{\boldsymbol{i}_0})\},$$

$$\breve{\breve{N}}_{k+1}=\varphi_{k+1}(\alpha_{k+1}G_k(Z_k)(1+\epsilon'),V_{\boldsymbol{i}_0}).$$

- If $\alpha_{k+1}G_k(Z_k)\times(1+\epsilon')>1$, we set

$$\alpha'(Z_k)=2-\alpha_{k+1}G_k(Z_k)(1+\epsilon').$$

As $\alpha_{k+1}G_k(z)\leq 1$ for all z, we have

$$\alpha'(Z_k)\in[0,1],$$

We define a function $\breve{\varphi}_{k+1}$ such that

$$\breve{\varphi}_{k+1}(\alpha_{k+1}G_k(Z_k)(1+\epsilon'),V_{\boldsymbol{i}_0})=\begin{cases} i & \text{if } V_{\boldsymbol{i}_0}\in\left[\frac{i-1}{q_{k+1}},\frac{i-1}{q_{k+1}}+\frac{\alpha'(Z_k)}{q_{k+1}}\right), \\ 2i & \text{if } V_{\boldsymbol{i}_0}\in\left[\frac{i-1}{q_{k+1}}+\frac{\alpha'(Z_k)}{q_{k+1}},\frac{i}{q_{k+1}}\right). \end{cases}$$

And we set

$$Z^i_{k+1}=m_{k+1}(Z_k,U_{(\boldsymbol{i}_0,i)}),\ \text{for } i\in\{1,\dots,\breve{\varphi}_{k+1}(\alpha_{k+1}G_k(Z_k)(1+\epsilon'),V_{\boldsymbol{i}_0})\},$$

$$\breve{\breve{N}}_{k+1}=\breve{\varphi}_{k+1}(\alpha_{k+1}G_k(Z_k)(1+\epsilon'),V_{\boldsymbol{i}_0}).$$

We have

$$\begin{aligned}
\sum_{i=1}^{q_{k+1}} i\times\frac{\alpha'(Z_k)}{q_{k+1}}+2i\times\frac{1-\alpha'(Z_k)}{q_{k+1}} &= \frac{(q_{k+1}+1)}{2}(\alpha'(Z_k)+2(1-\alpha'(Z_k))) \\
&= \frac{(q_{k+1}+1)}{2}\alpha_{k+1}G_k(Z_k)(1+\epsilon').
\end{aligned}$$

We have the inequality

$$\varphi_{k+1}(\widetilde{G}_k(\boldsymbol{i}_0, Z_k, \Theta), V_{\boldsymbol{i}_0})\mathbb{1}_{|Z_k-Y_k|\leq s^2\epsilon/\delta} \leq \breve{\breve{N}}_{k+1}, \text{ a.s.}$$

So

$$\begin{aligned}
(5.11)\quad \mathbb{E}(\breve{N}_T(\Theta,k,Z_k)\mathbb{1}_{|Z_k-Y_k|\leq s^2\epsilon/\delta}) &\leq \mathbb{E}(\mathbb{1}_{|Z_k-Y_k|\leq s^2\epsilon/\delta}\sum_{i=1}^{\breve{\breve{N}}_{k+1}} \breve{N}_T(\Theta,k+1,Z_{k+1}^i)) \\
&= \mathbb{E}(\mathbb{1}_{|Z_k-Y_k|\leq s^2\epsilon/\delta}\breve{\breve{N}}_{k+1}\breve{N}_T(\Theta,k+1,Z_{k+1}^1)) \\
&= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\mathbb{1}_{|z-Y_k|\leq s^2\epsilon/\delta}\mathbb{1}_{\alpha_{k+1}G_k(z)(1+\epsilon')\leq 1}\left(\sum_{i=1}^{q_{k+1}} i\times\frac{\alpha_{k+1}G_k(z)(1+\epsilon')}{q_{k+1}}\right) \\
&\qquad \times \mathbb{E}(\breve{N}_T(\Theta,k+1,z'))\pi_k(dz)Q(z,dz') \\
&\quad + \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\mathbb{1}_{|z-Y_k|\leq s^2\epsilon/\delta}\mathbb{1}_{\alpha_{k+1}G_k(z)(1+\epsilon')>1}\left(\sum_{i=1}^{q_{k+1}} i\times\frac{\alpha'(z)}{q_{k+1}}+2i\times\frac{(1-\alpha'(z))}{q_{k+1}}\right) \\
&\qquad \times \mathbb{E}(\breve{N}_T(\Theta,k+1,z'))\pi_k(dz)Q(z,dz') \\
&= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\mathbb{1}_{|z-Y_k|\leq s^2\epsilon/\delta}\frac{\alpha_{k+1}(q_{k+1}+1)}{2}G_k(z)(1+\epsilon')\mathbb{E}(\breve{N}_T(\Theta,k+1,z'))\pi_k(dz)Q(z,dz') \\
&= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\mathbb{1}_{|z-Y_k|\leq s^2\epsilon/\delta}\frac{1}{\pi_k(G_k)}G_k(z)(1+\epsilon')\mathbb{E}(\breve{N}_T(\Theta,k+1,z'))\pi_k(dz)Q(z,dz') \\
&\leq (1+\epsilon')\mathbb{E}(\breve{N}_T(\Theta,k+1,Z_{k+1}))\,.
\end{aligned}$$

From Equations (5.10) and (5.11), we get

$$\begin{aligned}
\mathbb{E}(\breve{N}_T(\Theta,k,Z_k)) &\leq \frac{1}{\pi_k(G_k)}\frac{e^{-\frac{1}{2s^2}\left(\frac{s^2\epsilon}{\delta}-\sqrt{3}\delta\right)_+^2}}{(2\pi s^2)^{3/2}}\prod_{i=k+2}^{T} q_i + (1+\epsilon')\mathbb{E}(\breve{N}_T(\Theta,k+1,Z_{k+1}) \\
&\leq e^{-\frac{1}{2s^2}\left(\frac{s^2\epsilon}{\delta}-\sqrt{3}\delta\right)_+^2}\times\prod_{i=k+1}^{T} q_i + (1+\epsilon')\mathbb{E}(\breve{N}_T(\Theta,k+1,Z_{k+1})\,.
\end{aligned}$$

As $\breve{N}_T(\Theta,T,Z_T)=1$, we get by recurrence, for all k,

$$\mathbb{E}(\breve{N}_T(\Theta,k,Z_k)) \leq e^{-\frac{1}{2s^2}\left(\frac{s^2\epsilon}{\delta}-\sqrt{3}\delta\right)_+^2}\left(\sum_{i=k+1}^{T}(1+\epsilon')^{i-(k+1)}q_iq_{i+1}\dots q_T\right)+(1+\epsilon')^{T-k}\,.$$

For all k,

$$q_k \leq \frac{2\|G\|_\infty}{\pi_{k-1}(G_{k-1})}\,,$$

and (using Lemma 6.6)

$$\log\left(\frac{1}{\pi_k(G_k)}\right) = \frac{|Y_k-m_k|^2}{2s^2}\left(\frac{1/(2\sigma_\infty^2)}{1/(2s^2)+1/(2\sigma_\infty^2)}\right)-\frac{3}{2}\log\left(\frac{s^2+\sigma_\infty^2}{2\pi}\right)\,.$$

Which implies there exists a random variable K such that

$$\forall T\,,\ \sum_{i=2}^{T}\log(q_i)\leq KT, \text{ a.s.}$$

So, for all k, we have a.s.

$$\begin{aligned}
\mathbb{E}(\breve{N}_T(\Theta,k,Z_k)) &\leq e^{-\frac{1}{2s^2}\left(\frac{s^2\epsilon}{\delta}-\sqrt{3}\delta\right)_+^2}T\frac{((1+\epsilon')^{2T}+(q_2q_3\dots q_T)^2)}{2}+(1+\epsilon')^T \\
&\leq e^{-\frac{1}{2s^2}\left(\frac{s^2\epsilon}{\delta}-\sqrt{3}\delta\right)_+^2}T((1+\epsilon')^{2T}+e^{2KT})+(1+\epsilon')^T
\end{aligned}$$

$$\begin{aligned}(\text{for } T \text{ large enough}) \le \exp&\left(\frac{-1}{2s^2}\left(s^2\frac{C_\epsilon}{T}\left(\frac{C_\delta}{T^2}\wedge 1\right)^{-1}-\sqrt{3}\left(\frac{C_\delta}{T^2}\wedge 1\right)\right)_+^2\right)\\ &\times T\left(\exp\left(\sqrt{3}(e-1)C_\epsilon\right)+e^{2KT}\right)\\ &\qquad\qquad\qquad + \exp\left(\sqrt{3}(e-1)C_\epsilon\right) = O(1)\end{aligned}$$

□

Proof of Proposition 5.1. We fix $k \in [T-1]$. Let $x_k \in \mathbb{R}$. We set

$$x_{k+1}^j = m_{k+1}(x_k, U_{(\boldsymbol{B}(k),j)}) \text{ for } j \in \{2, \dots q_{k+1}\}\,.$$

We have

$$\widetilde{N}_T(\Theta, k, x_k) \le \sum_{j=2}^{q_{k+1}} \breve{N}_T(\Theta, k+1, x_{k+1}^j)\,.$$

Let $\delta > 0$. We fix j in $\{2, \dots, q_{k+1}\}$. We have, for any $x \in \mathbb{R}^3$ such that $V_{(\boldsymbol{B}(k),j)} > \alpha_{k+2}\widetilde{G}_k((\boldsymbol{B}(k), j), x, \Theta)$,

$$\mathbb{E}(\breve{N}_T(\Theta, k+1, Z_{k+1})|\Theta) \ge \delta^3 \times \inf_{y \in L_\delta(x)} \pi_{k+1}(y) \times \breve{N}_T(\Theta, k+1, x)\,.$$

So

$$\begin{aligned}\sup_x \breve{N}_T(\Theta, k+1, x) &\le \frac{\mathbb{E}(\breve{N}_T(\Theta, k+1, Z_{k+1})|\Theta)}{\delta^3 \times \inf\{\pi_{k+1}(y),\, y \in L_\delta(x),\, x : V_{(\boldsymbol{B}(k),j)} > \alpha_{k+2}\widetilde{G}_{k+1}((\boldsymbol{B}(k), j), x, \Theta)\}}\\ &\le \frac{\mathbb{E}(\breve{N}_T(\Theta, k+1, Z_{k+1})|\Theta)}{\delta^3 \times \inf\{\pi_{k+1}(y),\, y \in L_\delta(x),\, x : V_{(\boldsymbol{B}(k),j)} > \alpha_{k+2}G_{k+1}(x)\}}\,.\end{aligned}$$

We set

$$A = \mathbb{E}(\breve{N}_T(\Theta, k+1, Z_{k+1})|\Theta)\,,$$
$$B = \delta^3 \inf\{\pi_{k+1}(y),\, y \in L_\delta(x),\, x : V_{(\boldsymbol{B}(k),j)} > \alpha_{k+2}G_{k+1}(x)\}\,.$$

We are looking for a constant $\mu_{2,k+1}$ such that,

$$\mathbb{P}\left(\sup_x \breve{N}_T(\Theta, k+1, x) \ge \frac{4\mu_{2,k+1}}{q_{k+1}}\right) \le \frac{1}{4T}\,.$$

For this, it is sufficient to have (see Lemma 6.5)

$$\mathbb{P}\left(\frac{1}{B} \ge \frac{\mu_{2,k+1}}{2q_{k+1}T\mathbb{E}(A)}\right) \le \frac{1}{8T}\,.$$

By Lemma 5.2, we know that there exists almost surely a random variable C_f depending on $(Y_k)_{k\ge1}$ such that for all T, k,

$$\mathbb{E}(A) \le C_f\,.$$

So we are looking for $\mu_{2,k+1}$ such that

$$\mathbb{P}\left(\frac{1}{B} \ge \frac{\mu_{2,k+1}}{2C_fTq_{k+1}}\right) \le \frac{1}{8T}\,. \tag{5.12}$$

We have

$$\begin{aligned}&\delta^3 \inf\{\pi_{k+1}(y),\, y \in L_\delta(x),\, x : V_{(\boldsymbol{B}(k),j)} > \alpha_{k+2}G_{k+1}(x)\}\\ &\qquad\qquad \ge \frac{\delta^3}{(2\pi\sigma_\infty^2)^{3/2}} \exp\left(-\frac{1}{2\sigma_\infty^2}(\delta\sqrt{3} + |m_{k+1}^\pi - Y_{k+1}| + R)^2\right)\,,\end{aligned}$$

where $R \ge 0$ is such that

$$V_{(\boldsymbol{B}(k),j)} = \alpha_{k+2}G_{k+1}(R(1,1,1) + Y_{k+1})\,.$$

So

$$\alpha_{k+2}\frac{\exp\left(-\frac{1}{2s^2}R^2\right)}{(2\pi s^2)^{3/2}} \geq V_{(\boldsymbol{B}(k),j)}\,.$$

So

$$\begin{aligned} R &\leq \sqrt{-2s^2\log\left((2\pi s^2)^{3/2}\left(\frac{V_{(\boldsymbol{B}(k),j)}}{\alpha_{k+2}}\right)\right)} \\ (\text{as } \alpha_{k+2}\leq 1/\|G_{k+1}\|_\infty) &\leq \sqrt{-2s^2\log\left(V_{(\boldsymbol{B}(k),j)}\right)} \qquad (5.13) \end{aligned}$$

We have $m^\pi_{k+1}-Y_{k+1}\sim\mathcal{N}(0;(\sigma_\infty^2+s^2)\mathrm{Id}_3)$. So, to have Equation (5.12), is is sufficient to have

$$\mathbb{P}\left(\delta\sqrt{3}+|m^\pi_{k+1}-Y_{k+1}|+R\geq\sqrt{2\sigma_\infty^2\log\left(\frac{\delta^3}{(2\pi\sigma_\infty^2)^{3/2}}\frac{\mu_{2,k+1}}{2C_fTq_{k+1}}\right)}\right)\leq\frac{1}{8T}\,,$$

which is implied by (with Z of density $r\in\mathbb{R}\mapsto \mathbb{1}_{\mathbb{R}^+}(r)e^{-r^2/2(\sigma_\infty^2+s^2)}(2\pi(\sigma_\infty^2+s^2))^{-3/2}4\pi r^2$)

$$\begin{cases} \mathbb{P}\left(Z\geq\frac{1}{2}\left(\sqrt{2\sigma_\infty^2\log\left(\frac{\delta^3}{(2\pi\sigma_\infty^2)^{3/2}}\frac{\mu_{2,k+1}}{2C_fTq_{k+1}}\right)}-\delta\sqrt{3}\right)\right)\leq\frac{1}{16T}\,, \\ \mathbb{P}\left(R\geq\frac{1}{2}\left(\sqrt{2\sigma_\infty^2\log\left(\frac{\delta^3}{(2\pi\sigma_\infty^2)^{3/2}}\frac{\mu_{2,k+1}}{2C_fTq_{k+1}}\right)}-\delta\sqrt{3}\right)\right)\leq\frac{1}{16T}\,. \end{cases} \qquad (5.14)$$

We have for all $x>0$ (using integration by parts)

$$\begin{aligned} \mathbb{P}(Z\geq x) &\leq \sqrt{\frac{2}{\pi}}\left(\frac{x}{\sqrt{\sigma_\infty^2+s^2}}+\frac{\sqrt{\sigma_\infty^2+s^2}}{x}\right)e^{-x^2/(2(\sigma_\infty^2+s^2))} \\ (\text{if } x\geq 2s) &\leq \sqrt{\frac{2}{\pi}}\left(\frac{x}{\sqrt{\sigma_\infty^2+s^2}}+1\right)e^{-x^2/(2(\sigma_\infty^2+s^2))} \\ &\leq \sqrt{\frac{2}{\pi}}(e^{-x^2/(4(\sigma_\infty^2+s^2))}+e^{-x^2/(2(\sigma_\infty^2+s^2))}) \\ &\leq \sqrt{\frac{2}{\pi}}\times 2e^{-x^2/(4(\sigma_\infty^2+s^2))}\,. \end{aligned}$$

So, to have the first line of Equation (5.14), it is sufficient to have

$$\frac{1}{2}\left(\sqrt{2\sigma_\infty^2\log\left(\frac{\delta^3}{(2\pi\sigma_\infty^2)^{3/2}}\frac{\mu_{2,k+1}}{2C_fTq_{k+1}}\right)}-\delta\sqrt{3}\right)\geq 2\sqrt{\sigma_\infty^2+s^2}\vee\sqrt{(-4(\sigma_\infty^2+s^2))\log\left(\sqrt{\frac{\pi}{2}}\times\frac{1}{32T}\right)}\,,$$

that is

$$\begin{aligned} \mu_{2,k+1} \geq{}& \frac{2C_fTq_{k+1}(2\pi\sigma_\infty^2)^{3/2}}{\delta^3} \\ &\times\exp\left(\frac{1}{2\sigma_\infty^2}\left(2\left(2\sqrt{\sigma_\infty^2+s^2}\vee\sqrt{(-4(\sigma_\infty^2+s^2))\log\left(\sqrt{\frac{2}{\pi}}\times\frac{1}{32T}\right)}\right)+\delta\sqrt{3}\right)^2\right)\,. \end{aligned} \qquad (5.15)$$

By Equation (5.13), the second line of Equation (5.14) is implied by

$$\mathbb{P}\left(V_{(\boldsymbol{B}(k),j)}\leq\exp^{-\frac{1}{2s^2}\left(\frac{1}{2}\left(\sqrt{2\sigma_\infty^2\log\left(\frac{\delta^3}{(2\pi\sigma_\infty^2)^{3/2}}\frac{\mu_{2,k+1}}{2C_fTq_{k+1}}\right)}-\delta\sqrt{3}\right)\right)^2}\right)\leq\frac{1}{16T}\,,$$

which is equivalent to

$$\exp^{-\frac{1}{2s^2}\left(\frac{1}{2}\left(\sqrt{2\sigma_\infty^2\log\left(\frac{\delta^3}{(2\pi\sigma_\infty^2)^{3/2}}\frac{\mu_{2,k+1}}{2C_fTq_{k+1}}\right)}-\delta\sqrt{3}\right)\right)^2}\leq\frac{1}{16T}\,,$$

which is implied by

$$\mu_{2,k+1} \geq \frac{2C_f T q_{k+1}(2\pi\sigma_\infty^2)^{3/2}}{\delta^3} \exp\left(\frac{1}{2\sigma_\infty^2}\left(\delta\sqrt{3} + 2\left(-2s^2\log\left(\frac{1}{16T}\right)\right)^{\frac{1}{2}}\right)^2\right) . \tag{5.16}$$

We take $\mu_{2,k+1}$ to be the supremum of the right-hand sides of Equations (5.15), (5.16). And so there exists $C(\sigma, s)$ (a constant depending on σ and s) such that

$$\mu_{2,k+1} \leq \frac{C(\sigma, s)C_f q_{k+1}}{\delta^3} T^\gamma ,$$

with

$$\begin{aligned} \gamma &= 1 + \frac{4(\sigma_\infty^2 + s^2)}{\sigma_\infty^2} \vee \frac{2s^2}{\sigma_\infty^2} \\ &= 5 + \frac{4s^2}{\sigma_\infty^2} . \end{aligned} \tag{5.17}$$

So

$$\sum_{k=2}^{T} \mu_{2,k} \leq \frac{C(\sigma, s)C_f}{\delta^3} T^\gamma \textstyle\sum_{k=2}^{T} q_k .$$

Under Hypothesis 5, we have for all k in $[T-1]$,

$$q_{k+1} \leq \frac{2\|G_k\|_\infty}{\pi_k(G_k)} .$$

By Lemma 6.6, we have that

$$\frac{1}{\pi_k(G_k)} = \frac{(2\pi\sigma_\infty^2)^{3/2}(2\pi s^2)^{3/2}}{(2\pi\sigma_\infty^2 s^2/(\sigma_\infty^2 + s^2))^{3/2}} e^{|Y_k - m_k^\pi|^2 \frac{1}{2s^2}\left(\frac{1/(2\sigma_\infty^2)}{1/(2s^2)+1/(2\sigma_\infty^2)}\right)} .$$

Let us set

$$Y' = e^{|Y_k - m_k^\pi|^2 \frac{1}{2s^2}\left(\frac{1/(2\sigma_\infty^2)}{1/(2s^2)+1/(2\sigma_\infty^2)}\right)} .$$

It satisfies

$$\mathbb{P}(Y' \geq x) \underset{x\to+\infty}{\sim} \frac{2\sqrt{\log(x)}}{x\sqrt{\pi}} .$$

So, by Hypothesis 5 and adapting the Generalized Central-Limit Theorem of [UZ99] (p. 62),

$$\frac{1}{T^{\gamma+\iota}} \sum_{k=1}^{T-1} \mu_{2,k+1} \xrightarrow[T\to+\infty]{\text{law}} 0 ,$$

for all $\iota > 0$. □

Proposition 5.3. *Under Hypothesis 4, 5, if we choose $n_1(T)$ satisfying Equations (6.4) and (6.9) then*

$$\frac{1}{T^{\gamma'}}\left(\sum_{k=1}^{T-1} \mu_{2,k} + n_1 T\right) \underset{T\to+\infty}{\longrightarrow} 0 ,$$

with the same γ'as in Proposition 5.1.

By Remark 6.4, we can than say that the expectation of the complexity of the algorithm grows at most like a polynom in T. Remember that we are talkin here of the expectation conditional to the observations Y_1, Y_2, ...

Proof. By Lemma 4.2, we have, by recurrence,

$$\sigma_1(T)^2 \leq \sum_{k=1}^{T-1} [\pi_k(G_k)m_{2,k+1} + \pi_k(G_k^2)m_{1,k+1}^2 - 1] + 1 .$$

For all $k \in [T-1]$,

$$m_{1,k+1} = \frac{\alpha_{k+1}(q_{k+1}+1)}{2} = \frac{1}{\pi_k(G_k)}\,,$$

and

$$\begin{aligned} m_{2,k+1} &= \sum_{i=1}^{q_{k+1}} \frac{\alpha_{k+1} i^2}{q_{k+1}} \\ &= \frac{\alpha_{k+1}(q_{k+1}+1)(2q_{k+1}+1)}{6} \\ &= \frac{2q_{k+1}+6}{3\pi_k(G_k)} \\ \text{(using Equation (5.9))} \quad &\leq \frac{2(3\pi_k(G_k)^{-1}\|G_k\|_\infty - 1)+1}{3\pi_k(G_k)} \\ &\leq \frac{2\|G_k\|_\infty}{\pi_k(G_k)^2}\,. \end{aligned}$$

By Lemma 6.6, ($\propto$ meaning "proportional to")

$$\begin{aligned} \pi_k(G_k^2)m_{1,k+1}^2 &\propto \exp\left(|Y_k - m_k^\pi|^2\left(\frac{1}{s^2}\times\frac{(-1/(2\sigma_\infty^2))}{1/s^2+1/(2\sigma_\infty^2)} + \frac{2}{2s^2}\times\frac{1/(2\sigma_\infty^2)}{1/(2s^2)+1/(2\sigma_\infty^2)}\right)\right) \\ &= \exp\left(|Y_k - m_k^\pi|^2 \frac{1}{s^2}\frac{(-s^2(s^2+\sigma_\infty^2)+s^2(2\sigma_\infty^2+s^2)}{(2\sigma_\infty^2+s^2)(s^2+\sigma_\infty^2)}\right) \\ &= \exp\left(|Y_k - m_k^\pi|^2 \frac{\sigma_\infty^2}{2\sigma_\infty^2+s^2}\times\frac{1}{s^2+\sigma_\infty^2}\right)\,. \end{aligned}$$

As the variable $(Y_k - m_k^\pi)_{k\geq 1}$ are independent and of law $\mathcal{N}(0,(\sigma_\infty^2+s^2)\mathrm{Id}_3)$, we get that

$$\frac{1}{T}\sum_{k=1}^{T-1}\pi_k(G_k^2)m_{1,k+1}^2$$

converges almost surely when T goes to infinity.

As we have seen at the end of the proof of Proposition 5.1,

$$\frac{1}{T^{\gamma+\iota}}\sum_{k=1}^{T-1}\pi_k(G_k)m_{2,k+1} \xrightarrow[T\to+\infty]{\text{law}} 0$$

for all $\iota > 0$. The constant γ, defined in Equation (5.17) is bigger than 1. So

$$\frac{\sigma_1(T)^2}{T^{\gamma+\iota}} \xrightarrow[T\to+\infty]{\text{law}} 0\,,\ \forall \iota > 0\,.$$

We have supposed that $n_1(T)$ satisfies Equation (6.4) and (6.9). The expectation of the complexity of the algorithm is then given by Equation (6.10) (in Remark 6.4). As $\mu_1(T) = T$ by Lemma 4.1, we get the result. $\square$

5.2.4. *Complexity of the algorithm: a case study.* We suppose $A = 0.5 \times \mathrm{Id}_3$, the (V_n)'s follow the law $\mathcal{N}(0,(0.2)^2\mathrm{Id}_3)$, the $(W_n)'s$ follow the law $\mathcal{N}(0,(0.5)^2\mathrm{Id}_3)$, $\delta = 0.1$. We take here the laws of the number of children of an individual to be the same as in Subsection 4.1, Equation (4.2). We made a simulation of a sequence $(Y_k)_{1\leq k\leq T}$ and stored it. Then we calibrated the constants $(\alpha_{k+1}, q_{k+1})_{k\in[T-1]}$ (using Equations (4.7), (4.8)). In particular, these constants do not satisfy Hypothesis 5 but are supposed to be not far off.

Here the codes were written in `python` and are thus relatively slow to execute. Nevertheless, it takes a few minutes to sample a trajectory of length 50. Using Section 6.2, we decide to take

$$n_1(T) = \sup\left(\frac{16\sigma_1(T)^2}{\mu_1(T)}, \frac{\mu_2(T)}{\mu_1(T)}\right)\,.$$

From looking at Proposition 5.3, we hope that $n_1 = C(T)\times T$, with $C(T)$ growing as a polynom in T. We estimate $C(T) = n_1(T)/T$ for $T \in \{1, 2, \ldots, 20\}$ using Monte-Carlo (for each T, we used

T	5	10	15	20
$C(T)$	165.6	466.4	634.1	766.5

TABLE 2. $C(T)$ in a filtering case

10000 samples for the estimation of $\mu_1(T)$, $\sigma_1(T)^2$ defined in the Appendix, we used 100 samples for the estimation of each $\mathbb{E}(\widetilde{N}_T(\Theta,k))$ appearing in the definition of $\mu_2(T)$. We can then compare T and $C(T)$ see Table 2 for $T\in\{5,10,15,20\}$). A simple least square regression in log-log scale gives a slope of 1.21. So it seems sensible to take $n_1(T)$ proportional to T or $T^{3/2}$.

We now want to estimate the complexity of the whole algorithm. Due to Remark 6.4, this complexity is of the same order of the complexity of sampling a branching process and finding $\widetilde{N}_T(\Theta,k)$ for each k. Let us fix k. When we compute $\widetilde{N}_T(\Theta,k)$, we need to compute $\widetilde{N}_T(\Theta,x_k)$ for a finite number of x_k in a ball around Y_k (see equation (5.5)). This number is, in expectation, proportional to δ^{-3}. Taking $n_1=\lfloor T^{3/2}\rfloor$, the complexity of the algorithm for a fixed T is of order less than $T^{3/2}\mu_1(T)+\delta^{-3}\mu_2(T)$. We use the above estimates of $\mu_1(T)$, $\mu_2(T)$ for $T\in\{1,2,\dots,20\}$. We compare $T^{3/2}\mu_1(T)$ and $\mu_2(T)$ to T. A least square regression in log-log scale gives us a slope less than 2 in both cases. This means the complexity, as a function of T, grows like T^2. We know that the coefficient of proportionality between the complexity and T^2 includes δ^{-3}, so the algorithm is dimension dependent (it will be δ^{-d} in dimension d).

6. Appendix

6.1. **Convergence of branching process.** In this Section, we are interested in the branching process defined in Section 2.1. We suppose the following.

Hypothesis 6. *For all $k\geq 1$, $g\in[0,\|G_k\|_\infty]$,*

$$f_{k+1}(g,0)=1-\alpha_{k+1}g\,,$$

with $\alpha_{k+1}\in(0;1/\|G_k\|_\infty]$. For all $k\geq 1$, there exists $q_{k+1}\in\mathbb{N}^*$, $(p_{k+1,j})_{j\in[q_{k+1}]}\in\mathbb{R}_+^{q_{k+1}}$, such that, for all j in $[q_{k+1}]$, for all g in $[0,\|G_k\|_\infty]$,

$$f_{k+1}(g,j)=\alpha_{k+1}p_{k+1,j}g\,.$$

We define $m_{1,k+1}$ (for all $k\geq 1$) as in Equation (4.5).

Lemma 6.2 below shows that its empirical measure at time k converges to π_k (defined in Equation (4.3)).

We present here some results concerning the propagation of chaos for systems of particles. We use the results and notations of [DPR09]. Similar results can be found in [BPS91] (around p. 177).

For q and N in $\mathbb{N}^*$, we define

$$[N]^{[q]}=\{\text{applications from } [q] \text{ to } [N]\}\,,$$

$$\langle q,N\rangle=\begin{cases}\{\text{injections from } [q] \text{ to } [N]\} & \text{if } q\leq N\,,\\ \emptyset & \text{otherwise}\,.\end{cases}$$

For any sets E and F, we write

$$\{\text{injections from } E \text{ to } F\}=\{a:E\hookrightarrow F\}\,.$$

For an empirical measure

$$m(x_1,\dots,x_N)=\frac{1}{N}\sum_{i=1}^N\delta_{x_i}$$

(where $x_1,x_2,\dots,x_N$ in some set E), we define measures on E^q (for $q\in\mathbb{N}^*$, $N\in\mathbb{N}$) by their action on a test function φ,

$$m(x_1,\dots,x_N)^{\otimes q}(\varphi)=\begin{cases}\frac{1}{N^q}\sum_{a\in[N]^{[q]}}\varphi(x_{a(1)},\dots,x_{a(q)}) & \text{if } N>0\,.\\ 0 & \text{if } N=0\,.\end{cases}$$

$$m(x_1,\dots,x_N)^{\odot q}(\varphi)=\begin{cases}\frac{(N-q)!}{N!}\sum_{a\in\langle q,N\rangle}\varphi(x_{a(1)},\dots,x_{a(q)}) & \text{if } q\leq N\,,\\ 0 & \text{otherwise.}\end{cases}$$

The definition of $m(\dots)^{\otimes q}$ above is consistent with the usual tensor product of measures. For functions $\phi_1,\dots,\phi_q$, we define

$$\phi_1\otimes\dots\otimes\phi_q(x_1,\dots,x_q)=\phi_1(x_1)\phi_2(x_2)\dots\phi_q(x_q)\,.$$

We recall that for a sequence $(X_k)_{k\geq 0}$ of random variables in $\mathbb{R}$, and x in $\mathbb{R}$,

$$X_k\xrightarrow[k\to+\infty]{\mathbb{P}}x\Leftrightarrow X_k\xrightarrow[k\to+\infty]{\text{law}}x\,.$$

Lemma 6.1. *Suppose that $(N'_k)_{k\geq 1}$ are a random variables taking values in $\mathbb{N}$ such that*

$$N'_k\xrightarrow[k\to+\infty]{\mathbb{P}}+\infty\,.$$

Suppose we have random variables $(X^i_k)_{i,k\geq 1}$ taking values in a measurable space $(E,\mathcal{E})$. Suppose π is a probability measure on $(E,\mathcal{E})$. The two following points are equivalent.

(1) *For all bounded measurable φ,*

$$\frac{1}{N'_k}\sum_{i=1}^{N'_k}\varphi(X^i_k)\xrightarrow[k\to+\infty]{law}\pi(\varphi)\ \textit{(constant random variable)}\,.$$

(2) *For all $q\in\mathbb{N}^*$, $\phi_1,\dots,\phi_q$ bounded measurable from E to $\mathbb{R}^+$,*

$$\mathbb{E}(m(X^1_k,\dots,X^{N'_k}_k)^{\odot q}(\phi_1\otimes\dots\otimes\phi_q))\xrightarrow[k\to+\infty]{}\pi^{\otimes q}(\phi_1\otimes\dots\otimes\phi_q)\,.$$

The variables N'_k takes values in $\mathbb{N}$. For $k\in\mathbb{N}^*$ and $\omega\in\Omega$ such that $N'_k(\omega)=0$, we set, by convention,

$$\frac{1}{N'_k}\sum_{i=1}^{N'_k}\varphi(X^i_k)=0\,.$$

Proof. <u>(1)⇒(2)</u> We suppose we have (1). Corollary 4.3, p. 789 in [DPR09] tells us that for any q, n in $\mathbb{N}^*$ such that $q\leq n$ and any $x_1,\dots,x_n$ in E,

$$m(x_1,\dots,x_n)^{\odot q}=m(x_1,\dots,x_n)^{\otimes q}+\frac{1}{n}\overline{m}_q(x_1,\dots,x_n)\,,$$

where $\overline{m}_q$ is a signed measurable such that $\|\overline{m}_q\|_{\text{TV}}\leq B_q$, with a constant B_q independent of n. So, for any $q\in\mathbb{N}^*$, $\phi_1,\dots,\phi_q$ bounded measurable from E to $\mathbb{R}^+$,

$$\begin{aligned}\mathbb{E}(m(X^1_k,\dots,X^{N_k}_k)^{\odot q}(\phi_1\otimes\dots\otimes\phi_q))&=\mathbb{E}(m(X^1_k,\dots,X^{N_k}_k)^{\otimes q}(\phi_1\otimes\dots\otimes\phi_q)\\&\quad+\frac{1}{N_k}\overline{m}_q(X^1_k,\dots,X^{N_k}_k)(\phi_1\otimes\dots\otimes\phi_q))\end{aligned}$$

for some bounded measure $\overline{m}_q$. We have

$$\mathbb{E}(m(X^1_k,\dots,X^{N_k}_k)^{\otimes q}(\phi_1\otimes\dots\otimes\phi_q))=\mathbb{E}(\prod_{i=1}^q m(X^1_k,\dots,X^{N_k}_k)(\phi_i))\xrightarrow[k\to+\infty]{}\pi^{\otimes q}(\phi_1\otimes\dots\otimes\phi_q)\,.$$

So we have (2).

<u>(2)⇒(1)</u> We suppose we have (2). For any bounded measurable φ on E, we set

$$\overline{\varphi}:x\mapsto\varphi(x)-\pi(\varphi)\,.$$

We have

$$\mathbb{E}((m(X^1_k,\dots,X^{N_k}_k)-\pi(\varphi))^2)\quad=\quad\mathbb{E}\left(\frac{1}{N_k^2}\sum_{1\leq i\leq N_k}\sum_{1\leq j\leq N_k,j\neq i}\overline{\varphi}(X^i_k)\overline{\varphi}(X^j_k)\right)$$

$$+\mathbb{E}\left(\frac{1}{N_k^2}\sum_{1\leq i\leq N_k}\overline{\varphi}(X_k^i)^2\right).$$

We have

$$\mathbb{E}\left(\frac{1}{N_k^2}\sum_{1\leq i\leq N_k}\overline{\varphi}(X_k^i)^2\right)\underset{k\rightarrow+\infty}{\longrightarrow}0\,,$$

and

$$\begin{aligned}\mathbb{E}\left(\frac{1}{N_k^2}\sum_{1\leq i\leq N_k}\sum_{1\leq j\leq N_k, j\neq i}\overline{\varphi}(X_k^i)\overline{\varphi}(X_k^j)\right) &= \mathbb{E}\left(\frac{N_k(N_k-1)}{N_k^2}m(X_k^1,\dots,X_k^{N_k})^{\odot 2}(\overline{\varphi},\overline{\varphi})\right)\\ &\underset{k\rightarrow+\infty}{\longrightarrow} 0\,.\end{aligned}$$

So we have (1). □

Lemma 6.2. *Suppose that $\pi_k(G_k)>0$ for all k in $[T-1]$, then (under Hypothesis 6) for all k in $[T]$, all bounded measurable φ on E_k,*

$$N_k\underset{n_1\rightarrow+\infty}{\overset{\mathbb{P}}{\longrightarrow}}+\infty\,,\tag{6.1}$$

$$\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}\varphi(X_k^{\boldsymbol{i}})\underset{n_1\rightarrow+\infty}{\overset{law}{\longrightarrow}}\pi_k(\varphi)\ \textit{(constant random variable)}\,.\tag{6.2}$$

Proof. We prove it by recurrence. We use Lemma 6.1 many times in the following.

The number of particles at time 1 is $N_1=n_1$. The particles $X_1^1,\dots,X_1^{n_1}$ are i.i.d. of law M_1 (remember that $S_1=[n_1]$). So the law of large number gives us, for any bounded measurable φ on E_1,

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\varphi(X_{n_1}^i)\underset{n_1\rightarrow+\infty}{\overset{\text{a.s.}}{\longrightarrow}}M_1(\varphi)=\pi_1(\varphi)\,.$$

Suppose that Equations (6.1) and (6.2) are true for all j in $[k]$ ($k<T$). We have, for all bounded measurable φ,

$$\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}\varphi(X_k^{\boldsymbol{i}})\underset{n_1\rightarrow+\infty}{\overset{\text{law}}{\longrightarrow}}\pi_k(\varphi)\,,\ N_k\underset{n_1\rightarrow+\infty}{\overset{\mathbb{P}}{\longrightarrow}}+\infty\,.\tag{6.3}$$

Let us set $\mathcal{F}_k=\sigma(S_k,(X_k^{\boldsymbol{i}})_{\boldsymbol{i}\in S_k})$. For all $\boldsymbol{i}$ in S_k, we have

$$\mathbb{E}(N_{k+1}^{\boldsymbol{i}}|\mathcal{F}_k)\leq\mathbb{P}(N_{k+1}^{\boldsymbol{i}}\geq 1|\mathcal{F}_k)\times q_{k+1}\,.$$

For any M in $\mathbb{N}^*$,

$$\mathbb{1}_{N_k\geq M}\mathbb{P}\left(\sum_{\boldsymbol{i}\in S_k}N_{k+1}^{\boldsymbol{i}}\leq M|\mathcal{F}_k\right)\leq\inf\left(1,\sum_{S\subset S_k,\#S=(N_k-M)}\mathbb{P}(N_{k+1}^{\boldsymbol{i}}=0\,,\forall i\in S|\mathcal{F}_k)\right)\,,$$

and

$$\begin{aligned}\sum_{S\subset S_k,\#S=N_k-M}\mathbb{P}(N_{k+1}^{\boldsymbol{i}}=0\,,\forall i\in S|\mathcal{F}_k) &\leq \sum_{S\subset S_k,\#S=(N_k-M)}\prod_{\boldsymbol{i}\in S}\left(1-\frac{\mathbb{E}(N_{k+1}^{\boldsymbol{i}}|\mathcal{F}_k)}{q_{k+1}}\right)\\ &\leq\sum_{S\subset S_k,\#S=(N_k-M)}\exp\left(-\sum_{\boldsymbol{i}\in S}\frac{\mathbb{E}(N_{k+1}^{\boldsymbol{i}}|\mathcal{F}_k)}{q_{k+1}}\right)\\ &\leq\binom{N_k}{N_k-M}\exp\left(-N_k\times\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}m_{1,k+1}G_k(X_k^{\boldsymbol{i}})+M\right)\,.\end{aligned}$$

So

$$\mathbb{P}\left(N_{k+1}\leq M\right)\leq\mathbb{P}(N_k\leq M)$$
$$+\mathbb{E}\left(\inf\left(1,\binom{N_k}{N_k-M}\exp\left(-N_k\times\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}m_{1,k+1}G_k(X_k^{\boldsymbol{i}})+M\right)\right)\right).$$

Using (6.3), we see that

$$N_{k+1}\xrightarrow[n_1\to+\infty]{\mathbb{P}}+\infty\,.$$

For any $\boldsymbol{i}\in S_k$, knowing $X_k^{\boldsymbol{i}}$, $N_{k+1}^{\boldsymbol{i}}$ is of law $f(G_k(X_k^{\boldsymbol{i}}),.)$ and is independent of all $X_k^{\boldsymbol{j}}$, $N_{k+1}^{\boldsymbol{j}}$ for $\boldsymbol{j}\neq\boldsymbol{i}$. For any $q\in\mathbb{N}^*$, $\phi_1,\dots,\phi_q$ bounded measurable, from $\mathbb{N}\times E_k$ to $\mathbb{R}^+$,

$$\begin{aligned}
&\mathbb{E}\left(m((N_{k+1}^{\boldsymbol{i}},X_k^{\boldsymbol{i}})_{\boldsymbol{i}\in S_k})^{\odot q}(\phi_1\otimes\dots\otimes\phi_q)\right)\\
&\quad=\mathbb{E}\left(\mathbb{1}_{N_k\geq q}\frac{(N_k-q)!}{N_k!}\sum_{a:[q]\hookrightarrow S_k}\phi_1(N_{k+1}^{a(1)},X_k^{a(1)})\dots\phi_q(N_{k+1}^{a(q)},X_k^{a(q)})\right)\\
&\quad=\mathbb{E}\left(\mathbb{1}_{N_k\geq q}\frac{(N_k-q)!}{N_k!}\sum_{a:[q]\hookrightarrow S_k}\overline{\phi}_1(X_k^{a(1)})\dots\overline{\phi}_q(X_k^{a(q)})\right)\\
&\quad=\mathbb{E}(m((X_k^{\boldsymbol{i}})_{\boldsymbol{i}\in S_k})^{\odot q}(\overline{\phi}_1,\dots,\overline{\phi}_q))\xrightarrow[n_1\to+\infty]{}\prod_{j=1}^{q}\pi_k(\overline{\phi}_j)\,,
\end{aligned}$$

where

$$\overline{\phi}_i(x)=\sum_{r=1}^{q_{k+1}}\alpha_{k+1}p_{k+1,r}\phi_i(r,x)\,,\ \forall x\in E_k,\forall i\in[q]\,.$$

So, for any bounded measurable φ from E_k to $\mathbb{R}^+$,

$$\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}N_{k+1}^{\boldsymbol{i}}\varphi(X_k^{\boldsymbol{i}})\xrightarrow[n_1\to+\infty]{\text{law}}\pi_k(G_k\varphi)\times m_{1,k+1}\,.$$

So

$$\begin{aligned}
\mathbb{1}_{N_{k+1}\neq0}\frac{1}{N_{k+1}}\sum_{\boldsymbol{i}\in S_k}N_{k+1}^{\boldsymbol{i}}\varphi(X_k^{\boldsymbol{i}}) &= \mathbb{1}_{N_{k+1}\neq0}\frac{N_k}{\sum_{\boldsymbol{i}\in S_k}N_{k+1}^{\boldsymbol{i}}}\times\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}N_{k+1}^{\boldsymbol{i}}\varphi(X_k^{\boldsymbol{i}})\\
&\xrightarrow[n_1\to+\infty]{\text{law}} \frac{\pi_k(G_k\varphi)}{\pi_k(G_k)}\,.
\end{aligned}$$

Let us set

$$\widehat{\pi}_k(\varphi)=\frac{\pi_k(G_k\varphi)}{\pi_k(G_k)}\,,$$
$$\text{for all }\boldsymbol{j}\in S_{k+1}\,,\ \widehat{X}_{k+1}^{\boldsymbol{j}}=X_k^{\boldsymbol{j}(k)}\,.$$

We have

$$\frac{1}{N_{k+1}}\sum_{\boldsymbol{j}\in S_{k+1}}\varphi(\widehat{X}_{k+1}^{\boldsymbol{j}})\xrightarrow[n_1\to+\infty]{\text{law}}\widehat{\pi}_k(\varphi)\,.$$

Knowing $(\widehat{X}_{k+1}^{\boldsymbol{j}})_{\boldsymbol{j}\in S_{k+1}}$, the $(X_{k+1}^{\boldsymbol{j}})_{\boldsymbol{j}\in S_{k+1}}$ are independent and for all $\boldsymbol{j}\in S_{k+1}$,

$$X_{k+1}^{\boldsymbol{j}}\sim M_{k+1}(\widehat{X}_k^{\boldsymbol{j}(k)},.)\,.$$

So, for any $q\in\mathbb{N}^*$, $\phi_1,\dots,\phi_q$ bounded measurable, from $\mathbb{N}\times E_k$ to $\mathbb{R}^+$,

$$\begin{aligned}
\mathbb{E}\left(m((X_{k+1}^{\boldsymbol{j}})_{\boldsymbol{j}\in S_{k+1}})^{\odot q}(\phi_1\otimes\dots\otimes\phi_q)\right) &= \mathbb{E}\left(m((\widehat{X}_k^{\boldsymbol{j}(k)})_{\boldsymbol{j}\in S_{k+1}})^{\odot q}(M_{k+1}\phi_1\otimes\dots\otimes M_{k+1}\phi_q)\right)\\
&\xrightarrow[n_1\to+\infty]{} \widehat{\pi}_kM_{k+1}\otimes\dots\otimes\widehat{\pi}_kM_{k+1}(\phi_1\otimes\dots\otimes\phi_q)\\
&= \pi_{k+1}\otimes\dots\otimes\pi_{k+1}(\phi_1\otimes\dots\otimes\phi_q)\,,
\end{aligned}$$

where we use the notations

$$M_{k+1}\phi_1(x) = \int_{y\in E_{k+1}} M_{k+1}(x,dy)\phi_1(y)\,,\ \forall x \in E_k\,,$$

$$\widehat{\pi}_k M_{k+1}(dy) = \int_{x\in E_k} \widehat{\pi}_k(x) M_{k+1}(x,dy)\,(\text{measure on } E_{k+1}).$$

This finishes the proof. □

For all $k \in [T]$, we define $\mathcal{B}_1(E_k)$ to be the set of measurable functions $\varphi : E_k \to \mathbb{R}$ such that $\|\varphi\|_\infty \leq 1$.

Lemma 6.3. *We suppose here that the parameters $(\alpha_{k+1}, q_{k+1})_{k\in[T-1]}$ are chosen such that Equation (4.6) is true for all k in $[T-1]$. Then, for all φ in $\mathcal{B}_1(E_k)$, for all k in $[T]$,*

$$\mathbb{E}\left(\sum_{\boldsymbol{j}\in S_k\,:\,1\prec \boldsymbol{j}} \varphi(X_k^{\boldsymbol{j}})\right) = \pi_k(\varphi)\,.$$

Proof. We take k in $[T]$ and φ in $\mathcal{B}_1(E_k)$. We set

$$D_{k,n_1}(\varphi) = \left|\frac{1}{N_k}\sum_{\boldsymbol{i}\in S_k}\varphi(X_k^{\boldsymbol{i}}) - \pi_k(\varphi)\right|\,.$$

We have, for all $\epsilon > 0$,

$$\begin{aligned}
\left|\mathbb{E}\left(\frac{1}{n_1}\sum_{\boldsymbol{j}\in S_k}\varphi(X_k^{\boldsymbol{j}})\right) - \mathbb{E}\left(\frac{N_k}{n_1}\pi_k(\varphi)\right)\right| &\leq \mathbb{E}\left(\frac{N_k}{n_1}\epsilon\right) + \mathbb{E}\left(\frac{N_k}{n_1}\times 2\mathbb{1}_{[\epsilon,+\infty)}(D_{k,n_1}(\varphi))\right) \\
\text{(by Lemma 4.1)} &= \epsilon + \mathbb{E}\left(\frac{N_k}{n_1}\times 2\mathbb{1}_{[\epsilon,+\infty)}(D_{k,n_1}(\varphi))\right) \\
\text{(using Remark 4.3)} &= \epsilon + \mathbb{E}(\#\{\boldsymbol{i}\in S_k : 1\prec \boldsymbol{i}\}\times 2\mathbb{1}_{[\epsilon,+\infty)}(D_{k,n_1}(\varphi)))\,.
\end{aligned}$$

As $D_{k,n_1}(\varphi)$ converges to 0 in probability when n_1 converges to $+\infty$ (Lemma 6.2), we can find extract a subsequence, indexed by $(u(n_1))_{n_1\geq 1}$, such that

$$\mathbb{1}_{[\epsilon,+\infty)}(D_{k,u(n_1)}(\varphi)) \xrightarrow[n_1\to+\infty]{\text{a.s.}} 0\,.$$

As $\#\{\boldsymbol{i}\in S_k : 1\prec \boldsymbol{i}\}$ does not depend on n_1, we get (by dominated convergence)

$$\epsilon + \mathbb{E}(\#\{\boldsymbol{i}\in S_k : 1\prec \boldsymbol{i}\}\times 2\mathbb{1}_{[\epsilon,+\infty)}(D_{k,u(n_1)}(\varphi))) \xrightarrow[n_1\to+\infty]{} \epsilon\,.$$

We have (by Lemma 4.1 and Remark 4.3)

$$\mathbb{E}\left(\frac{N_k}{n_1}\pi_k(\varphi)\right) = \pi_k(\varphi)\,,$$

$$\mathbb{E}\left(\frac{1}{n_1}\sum_{\boldsymbol{j}\in S_k}\varphi(X_k^{\boldsymbol{j}})\right) = \mathbb{E}\left(\sum_{\boldsymbol{j}\in S_k\,:\,1\prec\boldsymbol{j}}\varphi(X_k^{\boldsymbol{j}})\right)\,.$$

So, for any $\epsilon > 0$,

$$\left|\mathbb{E}\left(\sum_{\boldsymbol{j}\in S_k\,:\,1\prec\boldsymbol{j}}\varphi(X_k^{\boldsymbol{j}})\right) - \pi_k(\varphi)\right| \leq \epsilon\,.$$

□

6.2. **Number of particles.**

6.2.1. *Choice of* n_1. We show here how to choose n_1 as a function of T. It applies to both the examples of Section 5.1 and 5.2. We set

$$\forall k \in [T]\,,\ \mu_1(k) = \mathbb{E}(\#\{\boldsymbol{i} \in S_k\ :\ 1 \prec \boldsymbol{i}\})\,,$$
$$\forall k \in [T]\,,\ \sigma_1(k)^2 = \mathbb{V}(\#\{\boldsymbol{i} \in S_k\ :\ 1 \prec \boldsymbol{i}\})$$

(the definition of σ_1 coincides with the one given in Section 2.1). We suppose we have constants

$$\mu_{2,k} \geq 1/4 \text{ such that } \forall k \in [T-1]\,,\ \mathbb{P}(\widetilde{\widetilde{N}}_T(\Theta, k) \geq 4\mu_{2,k}) \leq \frac{1}{4T}.$$

When the following expectation is finite, one can simply take, for all k,

$$\mu_{2,k} = T \times \mathbb{E}(\widetilde{\widetilde{N}}_T(\Theta, k))\,.$$

We then choose n_1 such that

$$n_1 - 1 \geq \frac{16\sigma_1(T)^2}{\mu_1(T)}\,. \tag{6.4}$$

It implies that (remember that, for any n_1, $\overline{N}_T(\Theta)$ is a sum of n_1 i.i.d. variables of mean $\mu_1(T)$ and variance $\sigma_1^2(T)$)

$$\begin{aligned}\mathbb{P}\left(\overline{N}_T(\Theta) \leq \frac{n_1\mu_1(T)}{2}\right) &= \mathbb{P}\left(\overline{N}_T(\Theta) - n_1\mu_1(T) \leq -\frac{n_1\mu_1(T)}{2}\right)\\ &\leq \frac{4\sigma_1^2(T)}{n_1\mu_1^2(T)} \leq \frac{1}{4}\,,\end{aligned} \tag{6.5}$$

and (remember that $\#\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \nexists k : \boldsymbol{B}(k) \prec \boldsymbol{i}\}$ is a sum of $n_1 - 1$ i.i.d. variables)

$$\mathbb{P}\left(\#\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \nexists k : \boldsymbol{B}(k) \prec \boldsymbol{i}\} \geq 2\mu_1(T)(n_1 - 1)\right) \leq \frac{\sigma_1^2(T)}{(n_1 - 1)\mu_1^2(T)} \leq \frac{1}{4}\,. \tag{6.6}$$

We have

$$\mathbb{P}\left(\sum_{k=1}^{T-1} \widetilde{\widetilde{N}}_T(\Theta, k) \geq 4\sum_{k=1}^{T-1}\mu_{2,k}\right) \leq \frac{1}{4}\,. \tag{6.7}$$

We set

$$\mu_2 = 1 + \sum_{k=1}^{T-1}\mu_{2,k}\,, \tag{6.8}$$

we write $\mu_2(T)$ instead of μ_2 when we want to stress the dependency in T. We choose n_1 such that it also satisfies

$$n_1\mu_1(T) \geq \mu_2(T)\,. \tag{6.9}$$

So, $\overline{N}_T(\Theta) \geq \frac{n_1\mu_1(T)}{2}$ and $\#\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \nexists k : \boldsymbol{B}(k) \prec \boldsymbol{i}\} \leq 2\mu_1(T)(n_1 - 1)$ and $1 + \sum_{k=1}^{T-1} \widetilde{\widetilde{N}}_T(\Theta, k) \leq 1 + 4\sum_{k=1}^{T-1}\mu_{2,k}$ implies (recall Equations (5.3), (5.6))

$$\begin{aligned}\frac{\overline{N}_T(\Theta)}{\widetilde{\widetilde{N}}_T(\Theta)} &\geq \frac{\left(\frac{n_1\mu_1(T)}{2}\right)}{2\mu_1(T)(n_1 - 1) + 1 + 4\sum_{k=1}^{T-1}\mu_{2,k}}\\ &\geq \frac{1}{12}\,.\end{aligned}$$

We have

$$\begin{aligned}\mathbb{P}\Bigg(\overline{N}_T(\Theta) \geq \frac{n_1\mu_1(T)}{2}, \#\{\boldsymbol{i} \in \widetilde{S}_T(\Theta_n), \nexists k : \boldsymbol{B}(k) \prec \boldsymbol{i}\} \leq 2\mu_1(T)(n_1 - 1)\,,\\ \sum_{k=1}^{T-1} \widetilde{\widetilde{N}}_T(\Theta, k) \leq 4\sum_{k=1}^{T-1}\mu_{2,k}\Bigg)\end{aligned}$$

$$\geq 1-\mathbb{P}\left(\overline{N}_T(\Theta)\leq\frac{n_1\mu_1(T)}{2}\right)-\mathbb{P}\left(\#\{\boldsymbol{i}\in\widetilde{S}_T(\Theta_n),\nexists k:\boldsymbol{B}(k)\prec\boldsymbol{i}\}\geq 2\mu_1(T)(n_1-1)\right)$$
$$-\mathbb{P}\left(\sum_{k=1}^{T-1}\widetilde{\widetilde{N}}_T(\Theta,k)\geq 4\sum_{k=1}^{T-1}\mu_{2,k}\right),$$

and using (using (6.5), (6.6), (6.7)), we see this last quantity is bigger than $\frac{1}{4}$. So

$$\mathbb{P}\left(\frac{\overline{N}_T(\Theta)}{\widetilde{\widetilde{N}}_T(\Theta)}\geq\frac{1}{12}\right)\geq\frac{1}{4}.$$

Remark 6.4. This means that the expected number of steps in the "repeat" loop of Algorithm 1 is bounded independently of T, provided n_1 satisfies (6.4) and (6.9). Due to the way the parameters are chosen in Section 4.1, the complexity of our algorithm is of order

$$\sum_{k=1}^{T-1}\mu_{2,k}+n_1T \tag{6.10}$$

in expectation.

6.3. **Technical Lemmas.**

Lemma 6.5. *Let A, B be random variables in $\mathbb{R}^+$, $\mathbb{R}^{+*}$ respectively, such that A is L^1. For all M, μ in $\mathbb{R}^{+*}$, we have*

$$\mathbb{P}\left(\frac{A}{B}\geq 4\mu\right)\leq\mathbb{P}\left(\frac{1}{B}\geq\frac{4\mu}{M\mathbb{E}(A)}\right)+\frac{1}{M}.$$

Proof. We have

$$\begin{aligned}\mathbb{P}\left(\frac{A}{B}<4\mu\right)&\geq\mathbb{P}\left(\frac{A}{B}<4\mu\,,\,A\leq M\mathbb{E}(A)\right)\\&\geq\mathbb{P}\left(\frac{M\mathbb{E}(A)}{B}<4\mu\,,\,A\leq M\mathbb{E}(A)\right)\\&\geq 1-\mathbb{P}\left(\frac{M\mathbb{E}(A)}{B}\geq 4\mu\right)-\mathbb{P}\left(A\geq M\mathbb{E}(A)\right),\end{aligned}$$

so

$$\mathbb{P}\left(\frac{A}{B}\geq 4\mu\right)\leq\mathbb{P}\left(\frac{1}{B}\geq\frac{4\mu}{M\mathbb{E}(A)}\right)+\mathbb{P}\left(A\geq M\mathbb{E}(A)\right)\leq\mathbb{P}\left(\frac{1}{B}\geq\frac{4\mu}{M\mathbb{E}(A)}\right)+\frac{1}{M}.$$

□

Lemma 6.6. *For the example of Section 5.2 and under the same assumptions as in Proposition 5.1, we have, for all $k\geq 2$,*

$$\pi_k(G_k)=\frac{(2\pi\sigma_\infty^2 s^2/(\sigma_\infty^2+s^2))^{3/2}}{(2\pi\sigma_\infty^2)^{3/2}(2\pi s^2)^{3/2}}e^{|Y_k-m_k|^2\frac{1}{2s^2}\left(\frac{1/(2s^2)}{1/(2s^2)+1/(2\sigma_\infty^2)}-1\right)}.$$

Proof. We compute

$$\begin{aligned}\pi_k(G_k)&=\int_{\mathbb{R}^3}\frac{e^{-|x|^2/(2\sigma_\infty^2)}}{(2\pi\sigma_\infty^2)^{3/2}}\frac{e^{-|x-(Y_k-m_k)|^2/(2s^2)}}{(2\pi s^2)^{3/2}}dx\\&=\frac{1}{(2\pi\sigma_\infty^2)^{3/2}(2\pi s^2)^{3/2}}\int_{\mathbb{R}^3}e^{-\left(\frac{1}{2\sigma_\infty^2}+\frac{1}{2s^2}\right)\left|x-(Y_k-m_k)\frac{1}{2s^2}\left(\frac{1}{2\sigma_\infty^2}+\frac{1}{2s^2}\right)^{-1}\right|^2}\\&\qquad\times e^{|Y_k-m_k|^2\frac{1}{2s^2}\left(\frac{1/(2s^2)}{1/(2s^2)+1/(2\sigma_\infty^2)}-1\right)}dx\\&=\frac{(2\pi\sigma_\infty^2 s^2/(\sigma_\infty^2+s^2))^{3/2}}{(2\pi\sigma_\infty^2)^{3/2}(2\pi s^2)^{3/2}}e^{|Y_k-m_k|^2\frac{1}{2s^2}\left(\frac{1/(2s^2)}{1/(2s^2)+1/(2\sigma_\infty^2)}-1\right)}.\end{aligned}$$

□

Acknowledgements

Arnaud Doucet's research was supported by the Engineering and Physical Sciences Research Council (grant EP/K000276/1, EP/K009850/1) and by the Air Force Office of Scientific Research/Asian Office of Aerospace Research and Development (AFOSR/AOARD) (grant AOARD-144042). Sylvain Rubenthaler's research benefited from a sabbatical in PIMS Vancouver (CNRS, UMI 3069) funded by the CNRS and from a month long professorship position in the Institute of Statistical Mathematics, Tokyo.

References

[ADH10] Christophe Andrieu, Arnaud Doucet, and Roman Holenstein, *Particle Markov chain Monte Carlo methods*, J. R. Stat. Soc. Ser. B Stat. Methodol. **72** (2010), no. 3, 269–342. MR 2758115

[BPS91] D. L. Burkholder, É. Pardoux, and A. Sznitman, *École d'Été de Probabilités de Saint-Flour XIX—1989*, Lecture Notes in Mathematics, vol. 1464, Springer-Verlag, Berlin, 1991, Papers from the school held in Saint-Flour, August 16–September 2, 1989, Edited by P. L. Hennequin. MR 1108182

[CT02] J. N. Corcoran and R. L. Tweedie, *Perfect sampling from independent Metropolis-Hastings chains*, J. Statist. Plann. Inference **104** (2002), no. 2, 297–314. MR 1906013

[DPR09] Pierre Del Moral, Frédéric Patras, and Sylvain Rubenthaler, *Tree based functional expansions for Feynman-Kac particle models*, Ann. Appl. Probab. **19** (2009), no. 2, 778–825. MR 2521888

[FT98] S. G. Foss and R. L. Tweedie, *Perfect simulation and backward coupling*, Comm. Statist. Stochastic Models **14** (1998), no. 1-2, 187–203, Special issue in honor of Marcel F. Neuts. MR 1617572 (99f:60123)

[Ken02] Tom Kennedy, *A faster implementation of the pivot algorithm for self-avoiding walks*, J. Statist. Phys. **106** (2002), no. 3-4, 407–429. MR 1884540

[Ken05] Wilfrid Kendall, *Notes on perfect simulation*, Markov chain Monte Carlo, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., vol. 7, World Sci. Publ., Hackensack, NJ, 2005, pp. 93–146. MR 2226849

[PW96] James Gary Propp and David Bruce Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995), vol. 9, 1996, pp. 223–252. MR 1611693 (99k:60176)

[UZ99] Vladimir V. Uchaikin and Vladimir M. Zolotarev, *Chance and stability*, Modern Probability and Statistics, VSP, Utrecht, 1999, Stable distributions and their applications, With a foreword by V. Yu. Korolev and Zolotarev. MR 1745764

C. Andrieu, Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom.
E-mail address: c.andrieu@bris.ac.uk

N. Chopin, C.R.E.S.T., Timbre J350 3, Avenue Pierre Larousse, 92240 MALAKOFF, France.
E-mail address: Nicolas.Chopin@ensae.fr

A. Doucet, Department of Statistics, 1 South Parks Road Oxford, OX1 3TG, United Kingdom.
E-mail address: doucet@stats.ox.ac.uk

S. Rubenthaler, Laboratoire J. A. Dieudonné, Université Nice Sophia Antipolis, Parc Valrose, 06108 Nice cedex 2, France.
E-mail address: rubentha@unice.fr