Binary branching processes with Moran type interactions

A. M. G. Cox[*,†], E. Horton[‡], D. Villemonais[§]

July 7, 2022

Abstract

The aim of this paper is to study the large population limit of a binary branching particle system with Moran type interactions: we introduce a new model where particles evolve, reproduce and die independently and, with a probability that may depend on the configuration of the whole system, the death of a particle may trigger the reproduction of another particle, while a branching event may trigger the death of an other one. We study the occupation measure of the new model, explicitly relating it to the Feynman-Kac semigroup of the underlying Markov evolution and quantifying the L^2 distance between their normalisations. This model extends the fixed size Moran type interacting particle system discussed in [18, 19, 6, 7, 57] and we will indeed show that our model outperforms the latter when used to approximate a birth and death process. We discuss several other applications of our model including the neutron transport equation [36, 15] and population size dynamics.

Keywords : interacting particle systems, branching processes, many-to-one, Markov processes, birth-and-death process, Moran model.

MSC: 82C22, 82C80, 65C05, 60J25, 92D25, 60J80.

1 Introduction

Branching processes naturally arise as pertinent models in population size dynamics [40, 41, 43], neutron transport [14], genetic dynamics [47], growth-fragmentation processes [4, 3] and cell proliferation kinetics [60], and as theoretical objects in their own right [40, 17, 37, 38, 39]. These models are characterised by the independence of the branching and killing events in the system, which leads to a multiplicative behaviour, as well as interest in the scaling behaviour and the characterisation of the system at large times. On the other hand, processes with Moran type interactions are natural models for finite populations with either variety-increasing or variety-reducing effects such as genetic drift, genetic mutations and natural selection. First

[*]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: a.m.g.cox@bath.ac.uk

[†]Supported by EPSRC Grant EP/P009220/1

[‡]Inria Bordeaux Research Center, 200 Avenue de la Vieille Tour, 33405 Talence, France. Email: emma.horton@inria.fr

[§]Institut Élie Cartan de Lorraine, Bureau 123, Université de Lorraine, 54506, Vandoeuvre-lès-Nancy Cedex, France. Email: denis.villemonais@univ-lorraine.fr

introduced by Moran [48], the Moran model describes the evolution of N genes such that at an exponential rate, two particles are chosen uniformly at random from the population, one of which is killed and the other splits in two. Thus, the independence between particles is lost. We refer the reader to [28] and references therein for an overview of this model and of its extensions. This type of resampling has since been employed in a range of particle system models to numerically solve Feynman-Kac models, see [12, 21, 20, 57] and references therein.

The model we consider in this paper provides a combination of these two types of dynamics: (natural) branching and killing, as well as Moran type interactions. More precisely, when the system is initiated from N particles, each particle evolves according to an independent copy of a given Markov process, X, until either a (binary) branching or killing event happens. Here, binary refers to the fact that the particle is replaced by exactly two independent copies of itself. If such a branching event occurs, with a probability that may depend on the configuration of the whole system, another particle is removed from the system a selection mechanism. Similarly, if a killing event occurs, with a probability which may also depend on the configuration of the whole system, another particle is duplicated via a resampling mechanism. We refer to this model as the *binary branching model with Moran interactions*, or BBMMI for short.

Our main result (cf. Theorem 6) gives an explicit relation between the average of the empirical distribution of the particle system and the average of the underlying Markov process X. Indeed, letting m_T denote the empirical distribution of the particle system at time T, we show that for any $T \geq 0$,

$$\mathbb{E}\left[\Pi_T^A \Pi_T^B m_T f\right] = \mathbf{E}_{m_0}\left[f(X_t)\exp\left(\int_0^t b(X_s) - \kappa(X_s)\,\mathrm{d}s\right)\mathbf{1}_{t<\tau_\partial}\right], \tag{1}$$

where b is the branching rate, κ is the soft killing rate, τ_∂ is the absorption time, Π_T^A and Π_T^B are weights that compensate for the resampling and selection events that occur up to time T, which we state explicitly in Theorem 6, and $\mathbf{E}_\cdot$ is the expectation with respect to the law of X. As we will later discuss, by choosing particular model parameters, we recover both the classical many-to-one formula [35] and the unbiased estimator proved in [57] for fixed size Moran type genetic algorithms. The second part of Theorem 6 then provides a precise bound on the L^2 distance between the normalised semigroup on the right-hand side of (1) and the normalised occupation measure of the particle system. In particular, it states that this distance converges to zero as the number of initial particles tends to infinity.

In the third section of this article, we illustrate several applications of this model by focussing on a particular selection/resampling mechanism in our model. More precisely, we consider a binary branching process whose population size is constrained to remain in $\{N_{min}, \dots, N_{max}\}$, where $2 \leq N_{min} \leq N_{max} < +\infty$. In order to constrain the size of the process, when the size of the population reaches N_{max} (resp. N_{min}) and a natural branching (resp. killing) event occurs, we set the probability of selection (resp. resampling) to be 1. We often refer to this model as the N_{min}–N_{max} model. As will be clear in the next section, the choice $N_{min} = 0$ and $N_{max} = +\infty$ is also allowed, so that the collection of models we consider ranges from constant population size models with neither branching nor killing (corresponding to the choice $N_{min} = N_{max}$) to branching population models (when $N_{min} = 0$ and $N_{max} = +\infty$) with or without a selection/resampling mechanism.

For comparable results concerning the resampling algorithm, we refer the reader to [29, 51, 13, 49] and to [27, 22, 45] for associated central limit theorems. Whether a central limit theorem can be proved in our setting remains open, but could be approached by carefully studying the square increments of the martingale decomposition used in the proof of Theorem 6.

Our model and results are also reminiscent of the genetic algorithms introduced by Del Moral (see [18, 19] and references therein), where a particle system with killing is constrained to remain at a constant size via a resampling mechanism. We also refer the reader to [6, 7, 32, 51, 56, 12] and references therein.

Finally, we note that our model fits into the more general class of controlled branching processes introduced by Sevastyanov and Zubkov in [52], where the number of reproductive individuals in each generation depends on the size of the previous generation via a control function. This work was later extended by Yanev [59] to allow for random control functions. We refer the reader to [52, 31, 55, 58] for discrete time versions of this process and to [30] for the continuous time version. In these articles, the authors study the convergence of the survival probability in different regimes, as well as the expected population size. We also note that the control function can be seen as a way to model immigration and emigration of particles, [58]. We refer the reader to [54, 42, 50, 46, 44, 2, 1] for results regarding these latter processes.

The rest of the paper is set out as follows. In the next section, we introduce some notation and assumptions, and formally describe the particle system introduced above. We will then state our main result in Theorem 6 and discuss several implications.

In section 3, we focus on the N_{min}–N_{max} model and consider a variety of applications in order to illustrate the scope of our results, as well as some of the possible difficulties and extensions. In particular, in Section 3.1, we describe and study a stochastic population model with constrained size, based on the N_{min}–N_{max} model. We give a sufficient condition ensuring that the number of resampling events does not explode in finite time when the killing rate of X is unbounded. We also relate the long-time convergence of the empirical distribution of the population to the existence of a quasi-stationary distribution for an auxiliary sub-Markov process related to X and show the time uniform convergence of the renormalised empirical distribution of the particle system toward its conditional distribution, at a polynomial speed when $\sqrt{N_{max}}/N_{min}$ goes to infinity. We prove these results in the particular setting of this model, but our methods could be extended to a more general setting.

In Section 3.2, we deal with the particular case where the underlying Markov process is a piecewise deterministic Markov process (PDMP). In this case, it is possible for two particles to be killed at the same time, meaning that our main result is not directly applicable without further work. Working in the particular setting of the neutron transport equation (NTE), see [36, 15], we overcome this difficulty using the notion of h-transform introduced in [15] for the NTE. We show that particles whose dynamics are given by an appropriately h-transformed process fulfil the assumptions of our main theorem, and thus, we may use the N_{min}–N_{max} model with the h-transformed process in order to obtain estimates for quantities associated to the original process.

In Section 3.3, we focus on some numerical properties of the resampling/selection process. First, we make comparisons between the N_{min}–N_{max} model and the fixed size Moran type interacting particle system (IPS) for the approximation of quantities related to non-conservative semigroups (cf. [18, 19, 6, 7]), which corresponds to the particular case $N_{min} = N_{max}$ with no branching. We observe that in the case of a birth-and-death process, X, with branching rate given by $X_t \wedge M$ for some $M > 0$ and state-dependent Markov transitions, the computational cost, measured as the number of resampling/selection events, is significantly lower when using the N_{min}–N_{max} algorithm. We also study the bias and standard deviation of an estimator for the normalised empirical stationary distribution for each of the N_{min}–N_{max} and fixed size Moran type IPS. Letting M go to infinity, we provide numerical evidence that the number of interactions, bias and standard deviation stabilise for the N_{min}–N_{max} process, whereas these quantities grow linearly for the fixed size Moran type IPS process. Based on these simulations, we conjecture that the results of Section 2.2 hold true in situations where the branching rate is unbounded and thus where the Moran type process is undefined. In the final part of Section 3.3, we introduce a filtering method to deal with some of the numerical discrepancies that occur when the variance of the number of resampling/selection events is high.

In Section 4 we provide the proof of Theorem 6. Finally, we end the paper with an appendix containing the formal construction of the particle system introduced in section 2.1.

2 Main results

2.1 Description of the model

Let $(\Omega, \mathcal{F}, (X_t)_{t\in[0,+\infty)})$ be a continuous time progressively measurable Markov process with values in a measurable state space $E \cup \partial$, where ∂ is an absorbing (measurable) set such that $\partial \cap E = \emptyset$. Denoting by $\tau_\partial = \inf\{t \geq 0,\ X_t \in \partial\}$ its absorption time, it follows that

$$X_t \in \partial,\ \forall t \geq \tau_\partial.$$

We denote by $\mathbf{P}_x$ its law when initiated at $x \in E$ and by $\mathbf{E}_x$ the corresponding expectation operator. We extend, whenever necessary, any measurable function $f : E \to [0, +\infty)$ by $f \equiv 0$ on ∂. We also assume that we are given two bounded functions $b : E \to \mathbb{R}_+$ and $\kappa : E \to \mathbb{R}_+$, denoting.

We will shortly define a particle system, where particles move according to copies of X between interactions. Before doing so, we make the following assumption, necessary for the system to be well defined. It entails that two independent copies of X are absorbed simultaneously with probability 0.

Assumption 1. For any $x \in E$ and $t \in [0, +\infty)$, $\mathbf{P}_x(\tau_\partial = t) = 0$ and $\mathbf{P}_x(\tau_\partial > t) > 0$.

We also introduce the following notation. Denote by $\mathcal{P}_f(\mathbb{N})$ the collection of finite subsets of $\mathbb{N} := \{1, 2, \dots\}$. Then, for $i \in \mathbb{N}$, we let

$$b_i : (E \cup \partial)^{\mathcal{P}_f(\mathbb{N})} \to [0, +\infty),$$
$$\kappa_i : (E \cup \partial)^{\mathcal{P}_f(\mathbb{N})} \to [0, +\infty)$$

be a collection of bounded measurable functions. Also for $i \in \mathbb{N}$, let

$$p_i : (E \cup \partial)^{\mathcal{P}_f(\mathbb{N})} \to [0, 1],$$
$$q_i : (E \cup \partial)^{\mathcal{P}_f(\mathbb{N})} \to [0, 1]$$

be measurable functions. We will assume that, for any $s \in \mathcal{P}_f(\mathbb{N})$, for any collection of points $\{x_i : i \in s, x_i \in E\}$, and any $i_0 \in s$,

$$b^{i_0}(x_i, i \in s) - \kappa^{i_0}(x_i, i \in s) = b(x_{i_0}) - \kappa(x_{i_0}) \tag{2}$$

and that

$$p^{i_0}(x_i, i \in s) = 0 \text{ whenever } |s| = 1, \tag{3}$$

where $|s|$ denotes the number of elements in the set s.

We now proceed with an informal description of the dynamic of the particle system, which we call the binary branching model with Moran type interactions (BBMMI). The formal construction of the process is a non-trivial task, and is given in the appendix. To this end, fix $\bar{N}_0 \geq 2$. We consider the particle system $(\bar{S}_t, (X^i_t)_{i\in\bar{S}_t})_{t\in[0,+\infty)}$, where the component $\bar{S}_t \in \mathcal{P}_f(\mathbb{N})$ is the set enumerating the particles in the system at time t, initiated with $\bar{S}_0 = \{1, 2, \dots, \bar{N}_0\}$ and $X^i_0 = x_i \in E$ for all $i \in \bar{S}_0$. The number of particles in the system at any time t is denoted by $\bar{N}_t = |\bar{S}_t|$.

Evolution of the BBMMI.

1. The particles X^i, $i \in \bar{S}_0$, evolve as independent copies of X, and we consider the following times:

$$\tau_1^{b,i} := \inf\{t \geq 0,\ \int_0^t b^i(X^i_s,\, i \in \bar{S}_0)\, ds \geq e^i_1\},$$

and

$$\tau_1^{\kappa,i} := \inf\{t \geq 0,\ \int_0^t \kappa^i(X^i_s,\, i \in \bar{S}_0)\, ds \geq E^i_1\}$$

and

$$\tau_1^{\partial,i} := \inf\{t \geq 0,\ X^i_t \in \partial\},$$

where e^i_1, E^i_1, $i = 1, \dots, \bar{N}_0$ are exponential random variables with parameter 1, and are independent of each other and everything else.

2. Denoting by i_0 the index of the (unique) particle such that $\tau_1^{b,i_0} \wedge \tau_1^{\kappa,i_0} \wedge \tau_1^{\partial,i_0} = \tau_1$, where $\tau_1 = \min_{i\in\bar{S}_0} \tau_1^{b,i} \wedge \tau_1^{\kappa,i} \wedge \tau_1^{\partial,i}$, we set $\bar{S}_t = \bar{S}_0$ for all $t < \tau_1$ and

   (a) if $\tau_1 = \tau_1^{b,i_0}$, then a *branching event* occurs;

   (b) if $\tau_1 = \tau_1^{\kappa,i}$, then a *soft killing event* occurs;

   (c) if $\tau_1 = \tau_1^{\partial,i_0}$, then a *hard killing event* occurs.

3. Then a resampling or selection event may occur, depending on the following situations:

   killing: if a (hard or soft) killing event occurred at the preceding step, then we say that i_0 is *killed* at time τ_1 and

   - with probability $p^{i_0}(X^i_{\tau_1},\, i \in \bar{S}_0)$, the particle i_0 is removed from the system and a *resampling event* occurs: one chooses j_0 uniformly in $S_0 \setminus \{i_0\}$ and sets

   $$X^{\max S_0+1}_{\tau_1} = X^{j_0}_{\tau_1} \text{ and } \bar{S}_{\tau_1} := \bar{S}_0 \setminus \{i_0\} \cup \{\max \bar{S}_0 + 1\};$$

     observe that the number of particles in the system at time τ_1 is then $\bar{N}_{\tau_1} = \bar{N}_0$; we say that j_0 is *duplicated* at time τ_1.
   - with probability $1-p^{i_0}(X^i_{\tau_1},\, i \in \bar{S}_0)$, the particle i_0 is removed from the system; the set of particles at time τ_1 is enumerated by $\bar{S}_{\tau_1} := \bar{S}_0 \setminus \{i_0\}$ and the number of particles in the system is then $\bar{N}_{\tau_1} = \bar{N}_0 - 1$;

   branching: if a branching event occurred at the preceding step, then we say that i_0 has *branched* at time τ_1 and

- with probability $q^{i_0}(X^i_{\tau_1}, i \in \bar{S}_0)$, a new particle is added to the system at position $X^{i_0}_{\tau_1}$ and a *selection event* occurs: one chooses j_0 at random uniformly in $S_0 \cup \{\max S_0 + 1\}$ and removes the particle j_0 from the system:

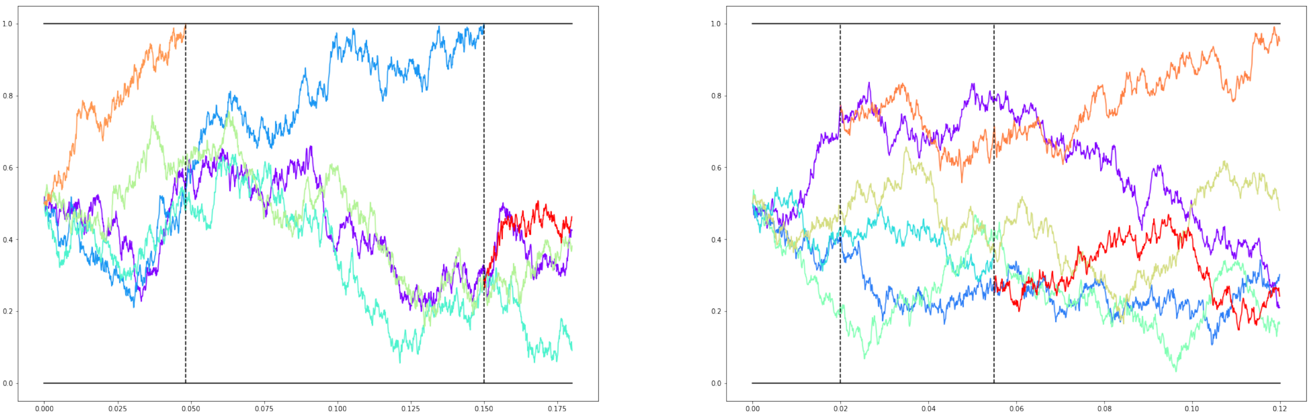
$$X^{\max S_0+1}_{\tau_1} = X^{i_0}_{\tau_1} \text{ and } \bar{S}_{\tau_1} := \{\max \bar{S}_0 + 1\} \cup \bar{S}_0 \setminus \{j_0\};$$

  observe that the number of particles at time τ_1 is thus $\bar{N}_{\tau_1} = \bar{N}_0$; we say that j_0 is *removed* at time τ_1 and that $\max S_0 + 1$ is *born* at time τ_1;
- with probability $1 - q^{i_0}(X^i_{\tau_1}, i \in \bar{S}_0)$, a new particle is added to the system at position $X^{i_0}_{\tau_1}$:

$$X^{\max S_0+1}_{\tau_1} = X^{i_0}_{\tau_1} \text{ and } \bar{S}_{\tau_1} = \bar{S}_0 \cup \{\max \bar{S}_0\},$$

  and we say that $\max S_0 + 1$ is *born* at time τ_1.

After time τ_1 the system evolves as independent copies of X until the next killing/branching event, denoted by τ_2, and at time τ_2 it may undergo a resampling/selection event as above. By iteration, we define the sequence $\tau_0 := 0 < \tau_1 < \tau_2 < \cdots < \tau_n < \cdots$.



(a) Resampling. Just before time 0.05, the orange particle is killed but no resampling occurs. Around time 0.15, the blue particle is killed and the green particle is resampled to produce the red particle.

(b) Selection. At time 0.02, the purple particle branches, producing the orange particle but no selection occurs. At the second dashed line, the dark blue particle branches to produce the red particle and the light blue particle is selected and removed from the system.

Figure 1: Simulations of the BBMMI model where X is a Brownian motion, killed at 0 and 1, with constant branching rate.

We will also make use of the following assumption, which ensures that the process described above is well defined at any time $t \geq 0$.

Assumption 2. The sequence $(\tau_n)_{n\in\mathbb{N}}$ converges to $+\infty$ almost surely.

Let us now make several remarks concerning the above particle system and potential parameter choices.

Remark 1. The enumeration of the particle system does not keep track of the genealogy of the particles in the system in order to simplify notation; this could be done using the Ulam tree as a set of indices.

Remark 2. The functions b^{i_0} and κ^{i_0} are the rate at which the particle i in the system branches and is killed respectively (in addition to hard killing events occurring when a particle hits the absorbing set ∂). The simplest situation is when $b^{i_0}(x^i,\, i\in s)=b(x_{i_0})$ and $\kappa^{i_0}(x_i,\, i\in s)=\kappa(x_{i_0})$, but other configurations are possible. We emphasise that the balance condition (2) is crucial to ensure the relation with the semigroup Q defined in the next section, and should not be considered a technical assumption.

Remark 3. The functions p^{i_0} (respectively q^{i_0}) denote the probability that a killing event (respectively a branching event) triggers a branching (respectively a killing) among the particles in the system. We call such an event a *resampling event* (respectively a *selection event*). In the context of Moran type models, they correspond to *selection by death* events (respectively *selection by birth* events). Condition (3) prevents a resampling event from occurring when the population decreases to 0 particles. We emphasize that, since ∂ is a set, p^{i_0} may depend on the killing location.

Remark 4. The above dynamic allows one to constrain the size of the process to remain between two bounds $N_{min}\neq 1$ and N_{max}, with $0\le N_{min}\le N_{max}$. In order to do so, one simply chooses

$$p^{i_0}(x_i,\, i\in s)=\mathbf{1}_{\{|s|=N_{min}\}} \text{ and } q^{i_0}(x_i,\, i\in s)=\mathbf{1}_{\{|s|=N_{max}\}}.$$

In what follows, we refer to this dynamic as the N_{min}–N_{max} process. Other natural choices include, for instance,

$$p^{i_0}(x_i,\, i\in s)=\frac{1}{\sharp s+1}\mathbf{1}_{|s|\ge 2} \text{ and } q^{i_0}(x_i,\, i\in s)=1-\frac{1}{\sharp s+1},$$

i.e. resamplings become less probable when the number of particles increases and selection becomes stronger when the number of particles increases, respectively. Note also that, since X is only assumed to be Markov, it can carry the time as a component, which would allow N_{min} and N_{max} to vary with time (in the former situation) or the selection pressure to vary with time (in the later example), by simply letting p,q depend on time.

Remark 5. In our model, we consider general killing, which corresponds to a conservative Markov process X killed at a possibly unbounded rate. Finding general criteria ensuring that Assumptions 1 and 2 hold is usually involved (it is trivial when the killing rate is bounded and there is no hard killing). In Section 3.1 we provide a criterion based on a Foster-Lyapunov type assumption ensuring the non-explosion of the system when the killing rate κ is unbounded (and when there is no hard killing) and also show in Section 3.2 that Assumption 2 holds true for a specific PDMP model, obtained as the h transform of a Neutron Random Walk killed at the boundary of domain. There is also a substantial literature on these types of problems in the presence of hard killing for the more classical fixed size Moran type systems, and, in particular, the interested reader may find it useful to adapt the methods developed in [7, 32, 56, 57] to our setting.

In the rest of this article, we denote by σ_n and ρ_n the n^{th} selection and resampling events, respectively. We let A_t denote the total number of resampling events up to time t, B_t denote the total number of selection events up to time t and we set $C_t=\inf\{n\ge 0 : t<\tau_{n+1}\}$ to be the total number of events up to time t. We will also use the following quantities throughout,

$$\Pi_t^A := \prod_{n=1}^{A_t}\left(\frac{\bar N_{\rho_n}-1}{\bar N_{\rho_n}}\right), \qquad \Pi_t^B := \prod_{i=1}^{B_t}\left(\frac{\bar N_{\sigma_n}+1}{\bar N_{\sigma_n}}\right). \tag{4}$$

2.2 L^2 bounds for the empirical distribution of the process

For all bounded measurable functions $f : E \to \mathbb{R}$, all $t \geq 0$ and all $x \in E$, we set

$$Q_t f(x) = \mathbf{E}_x \left[f(X_t) \exp\left(\int_0^t (b(X_s) - \kappa(X_s))\,\mathrm{d}s \right) \mathbf{1}_{t<\tau_\partial} \right],$$

where we recall that $\tau_\partial = \inf\{t \geq 0 : X_t \in \partial\}$. This defines a Feynman-Kac semigroup $(Q_t)_{t\geq 0}$, which is related to the binary branching model where particles move as copies of X that are killed at rate κ and branch at rate b resulting in the creation of two independent copies of the original particle. The relation between Q and this process is given by the well known many-to-one formula (see for instance [35] and references therein, see also Remark 8). In this section, we consider the BBMMI defined above and study its relation with Q.

In what follows, we let m_t (resp $\hat{m}_t$) denote the occupation measure (resp. normalised occupation measure) of the particle system,

$$m_t = \sum_{i\in\bar{S}_t} \delta_{X_t^i}, \qquad t \geq 0, \quad \hat{m}_t = \frac{1}{\bar{N}_0} m_t. \tag{5}$$

The following result first provides an equality between a penalised version of the occupation measure of the system at any time T and the semigroup Q at time T. The second part is dedicated to the control of the L^2 distance between renormalised versions of the semigroup Q and the occupation measure of the process.

Theorem 6. *Under Assumptions 1 and 2, the BBMMI satisfies, for all time $T \geq 0$ and all bounded measurable functions $f : E \to \mathbb{R}$, the following many-to-one formula*

$$m_0 Q_T f = \mathbb{E}\left(\Pi_T^A\, \Pi_T^B m_T f\right). \tag{6}$$

Moreover,

$$\left\| \frac{m_0 Q_T f}{m_0 Q_T \mathbf{1}_E} - \frac{m_T(f)}{m_T(\mathbf{1}_E)} \mathbf{1}_{m_T\neq 0} \right\|_2 \leq C\,\exp(c\|b\|_\infty T) \frac{\|f\|_\infty}{m_0 Q_T \mathbf{1}_E / \bar{N}_0} \frac{1}{\sqrt{\bar{N}_0}}, \tag{7}$$

where C, c are positive constants.

The proof of this result is given in Section 4 but let us first make some further comments on the model and the above result.

Remark 7. In Theorem 6, m_0 is assumed deterministic with total mass $\bar{N}_0$. The result immediately extends to the case where m_0 is random by taking the expectation in (6) and (7) (in the latter, the right hand term may be infinite). Moreover, the expressions may also be written in terms of the normalised measure, $\hat{m}$, in which case (6) becomes $\hat{m}_0 Q_T f = \mathbb{E}\left(\Pi_T^A\, \Pi_T^B \hat{m}_T f\right)$ and (7) can be written as

$$\left\| \frac{\hat{m}_0 Q_T f}{\hat{m}_0 Q_T \mathbf{1}_E} - \frac{\hat{m}_T(f)}{\hat{m}_T(\mathbf{1}_E)} \mathbf{1}_{\hat{m}_T\neq 0} \right\|_2 \leq C\,\exp(c\|b\|_\infty T) \frac{\|f\|_\infty}{\hat{m}_0 Q_T \mathbf{1}_E} \frac{1}{\sqrt{\bar{N}_0}},$$

making explicit the 'classical' dependence $\bar{N}_0^{-1/2}$ on the number of particles which appears on the right-hand side.

Remark 8. When $p = q \equiv 0$, the particle system has the dynamic of the classical binary branching process described just before the statement of the theorem. In this situation, we have $\Pi^A_T = \Pi^B_T = 1$ almost surely, and we thus recover the classical many-to-one formula.

Remark 9. Although we have remained in the setting of binary branching in this article, it should be possible to extend the definition of the BBMMI to allow for more general offspring distributions when a (natural) branching event occurs by modifying the selection mechanism. We intend to address this in future work.

Remark 10. The definition of our model allows a population with varying size by including branching, death and Moran type interactions, which are key features in the context of population dynamics. However, when restricted to the situation where $p = q \equiv 1$ (so that the size of the process remains constant equal to $\bar{N}_0$), Theorem 6 extends the results of [57] by allowing selection events. Similar results were also obtained in [51] (see also references therein and the recent paper [12] for an adaptation to discrete state spaces, with more general interaction mechanisms), where the process is constrained to remain of constant size $\bar{N}_0$ and with uniformly bounded killing rate. Our main contribution to this constant size setting is that our assumptions allow for hard killing events (a feature observed in several models, such as diffusion processes killed at the boundary of a domain), that we consider general Markov processes (allowing implicit dependence with respect to time and with respect to the past of the process), and do not require continuity of the branching/killing rates with respect to the empirical measure of the process. As is commonly observed in the literature, uniform convergence with respect to time can be obtained assuming additional mixing conditions on the semigroup Q (see the above references and also [26, 49, 53] for models with hard killing). We illustrate this property in Section 3.1.

Remark 11. It is also possible to treat branching processes with non-local branching in this setting. Indeed, consider a branching process where, between branching events, particles move according to a Markov process $(\xi_t)_{t\geq 0}$ from their point of creation, and when at position $x \in E$, at rate $\beta(x)$, the particle is removed from the system and replaced by two particles at (random) positions $x_1, x_2 \in E$, which may be different to x. Particles also undergo (soft) killing at rate $\kappa(x)$. Denoting by $(\psi_t)_{t\geq 0}$ the linear semigroup of this process, in this case, the many-to-one formula is given by

$$\psi_t[g](x) = \mathbf{E}_x\left[\mathrm{e}^{\int_0^t b(X_s)-\kappa(X_s)\mathrm{d}s} g(X_t)\mathbf{1}_{t<\tau_\partial}\right],$$

where $(X_t)_{t\geq 0}$ behaves like $(\xi_t)_{t\geq 0}$, except that at rate $2b(X_t)$, the process X_t jumps to a new location given by $(\delta_{x_1} + \delta_{x_2})/2$. We refer the reader to section 3.2 for an example of a branching process with non-local branching.

Remark 12. The main property of the process $(\bar{S}_t, (X^i_t)_{t\in\bar{S}_t})_{t\geq 0}$ that is used in the proof of Theorem 6 is the identity

$$\mathbb{E}^t_x\left[Q_{T-(t+s)}f(X^1_{t+s})\mathbf{1}_{t+s<\tau_1}\right] = \mathtt{E}^t\left[Q_{T-(t+s)}f(Y^1_{t+s})e^{-\int_t^{t+s} h_u\,\mathrm{d}u}\mathbf{1}_{t+s<\tau_\partial}\right]\mathbf{1}_{t<\tau_1}, \tag{8}$$

where $\mathtt{E}^t$ is the expectation on a probability space Ω' such that $(Y^i_u)_{u\geq t}$, $i \in \bar{S}_0$, are independent copies of X, starting from X^i_t at time t, and where

$$h_u := \sum_{j\in\bar{S}_0} b^j((Y^i_u)_{i\in\bar{S}_0}) + \kappa^j((Y^i_u)_{i\in\bar{S}_0}).$$

This identity is derived from the Markov property, and the assumption on the dynamics of the process $(\bar{S}_t, (X^i_t)_{t\in\bar{S}_t})_{t\geq 0}$. As such, it is possible to see that Theorem 6 still holds under milder conditions on the construction. Specifically, suppose that there exist a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and a progressively measurable process $(\bar{S}_t, (X^i_t)_{t\in\bar{S}_t})_{t\geq 0}$ defined on this space, such that the rates b^i, κ^i, p^i, q^i are no longer functions, but merely progressively measurable processes, satisfying the equivalents[1] of (2), (3):

$$b^i_t - \kappa^i_t = b(X^i_t) - \kappa(X^i_t)$$
$$p^i_t = 0 \quad \text{whenever } |\bar{S}_t| = 1.$$

Then the proof of Theorem 6 still holds, provided that condition (8) can be verified in the more general case, and where we interpret $\mathbb{E}^t_x$ as the expectation conditional on the full sigma-algebra $\mathcal{F}_t$. Of course, such a property is something that may, in certain settings, be obtained through construction.

An example where such a construction might be natural is in a distributed computing environment. In this case, one may look to break the particles into subsets which are handled by separate processors with minimal communication. In this case, one may choose to impose $N_{min} - N_{max}$-like criterion on each separate processor. Roughly, each processor might be allocated an initial set of particles, which evolve without communication (by only having killing/branching events), so long as the number of particles remains within $[N_{min}, N_{max}]$. Since global properties of the particle system are only needed when resampling or selection events occur, the need for communication events in the system, which cause computational bottlenecks, could be minimised by choosing to allow these events provided the number of particles on a given processor does not move outside a specified range. In such a case, the number of particles on a given processor would depend on the allocation of particles to given processors (and which may itself be randomised independently of the history of the particle system). Then the processes b^i and κ^i would not be simply functions of the particle positions (since we also need to know which particles are on the same processor), but would be expected to be $(\mathcal{F}_t)_{t\geq 0}$-progressively measurable, for an appropriate filtration, which may then be larger than the filtration generated by the particle histories.

3 Applications

3.1 Population size dynamics with constrained population size

In this section, we introduce a branching model for a population where the size is constrained between $N_{min} \geq 2$ and $N_{max} \geq N_{min}$ and demonstrate that the results from the previous section may be applied. In this model, we associate with each individual a health state $x \in E := \mathbb{Z}^d_+$, $d \geq 1$, which evolves according to the dynamics of a multi-dimensional birth and death process during the life of an individual. The parameters of the pure jump Markovian dynamics are denoted by

$$\begin{cases} \beta_i(x) & = \text{ transition rate from } x \text{ to } x + e_i, \\ \delta_i(x) & = \text{ transition rate from } x \text{ to } x - e_i, \end{cases}$$

[1] Technically, we need that the integrals of the respective processes are equal up to modification, that is, if we define $I^{i,1}_t := \int_0^t (b^i_s - \kappa^i_s)\,\mathrm{d}s$ and $I^{i,2}_t := \int_0^t (b^i(X^i_s) - \kappa(X^i_s))\,\mathrm{d}s$, then we require the processes $I^{i,1}$ and $I^{i,2}$ to agree up to a modification. Note that as a consequence of the difference $b(\cdot) - \kappa(\cdot)$ being bounded, this implies that the processes are also indistinguishable.

where $e_i \in \mathbb{Z}_+^d$ denotes the vector with 1 in the i^{th} position and 0 everywhere else, $\beta_i, \delta_i : \mathbb{Z}_+^d \to [0, +\infty)$ for all $i \in \{1, \dots, d\}$, and $\delta_i(x) = 0$ for all $x = (x_1, \dots, x_d)$ such that $x_i = 0$ (so that the process cannot leave the set $\mathbb{Z}_+^d$).

When its health state is $x \in \mathbb{Z}_+^d$, at rate $\kappa(x)$ an individual dies, where $\kappa : \mathbb{Z}_+^d \to [0, +\infty)$, and at rate $b(x)$, it gives birth to an identical individual with equal health state x, where $b : \mathbb{Z}_+^d \to [0, +\infty)$ is bounded. Note that, in our model, a higher health state may thus be associated with a worse health condition. This corresponds, for instance, to the situation where x is the size of a (harmful) parasite population or the concentration of a virus.

In addition, an external mechanism controls the total size of the individuals population: the total number of individuals is bounded above by N_{max} by removing an individual chosen randomly and uniformly in the population when the population size reaches $N_{max} + 1$, and the total number of individuals is lower bounded by N_{min} by cloning an individual chosen randomly and uniformly when the population size reaches $N_{min} - 1$ (in the situation where $N_{min} = 0$, this last mechanism does not operate).

Formally, the state space of our population model is $F := \bigcup_{N=N_{min}}^{N_{max}} [E]^N$, where $[E]^N$ denotes the unordered N-tuples of E. We represent the elements of F by $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$ and note that repetitions are allowed. The extended infinitesimal generator of the process acting on bounded mesurable functions $f : F \to \mathbb{R}$, is given, for any $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in F$, by

$$\begin{aligned}
\bar{L}f(\mathbf{x}) = \sum_{k=1}^{N} \sum_{i=1}^{d} \Big(\beta_i(\mathbf{x}_{k,i}) \Delta_{\mathbf{x}, \mathbf{x}+e_{k,i}^N} f + \delta_i(\mathbf{x}_{k,i}) \Delta_{\mathbf{x}, \mathbf{x}-e_{k,i}^N} f \Big) \\
+ \sum_{k=1}^{N} b(\mathbf{x}_k) \Delta_{\mathbf{x}, \mathbf{x} \cup [\mathbf{x}_k]} f + \sum_{k=1}^{N} \kappa(\mathbf{x}_k) \Delta_{\mathbf{x}, \mathbf{x} \setminus [\mathbf{x}_k]} f(\mathbf{x}) \\
+ \mathbf{1}_{N=N_{min}} \sum_{k=1}^{N} \kappa(\mathbf{x}_k) \frac{1}{N-1} \sum_{\ell=1, \ell \neq k}^{N} \Delta_{\mathbf{x} \setminus [\mathbf{x}_k], \mathbf{x} \cup [\mathbf{x}_\ell] \setminus [\mathbf{x}_k]} f \\
+ \mathbf{1}_{N=N_{max}} \sum_{k=1}^{N} b(\mathbf{x}_k) \frac{1}{N+1} \sum_{\ell=1}^{N} (1 + \mathbf{1}_{\ell=k}) \Delta_{\mathbf{x} \cup [\mathbf{x}_k], \mathbf{x} \cup [\mathbf{x}_k] \setminus [\mathbf{x}_\ell]} f,
\end{aligned}$$

where

- $\mathbf{x}_{k,i}$ is the i^{th} component of $\mathbf{x}_k$,
- $\mathbf{x} + e_{k,i}^N$ is $\mathbf{x}$ but with $\mathbf{x}_{k,i}$ replaced by $\mathbf{x}_{k,i} + 1$,
- $\mathbf{x} - e_{k,i}^N$ is $\mathbf{x}$ but with $\mathbf{x}_{k,i}$ replaced by $\mathbf{x}_{k,i} - 1$,
- for all $\mathbf{x}, \mathbf{y}$, $\Delta_{\mathbf{x},\mathbf{y}} f := f(\mathbf{y}) \mathbf{1}_{\mathbf{y} \in F} - f(\mathbf{x}) \mathbf{1}_{\mathbf{x} \in F}$,
- for all $y \in E$, $\mathbf{x} \cup [y] = [\mathbf{x}_1, \dots, \mathbf{x}_N, y]$,
- $\mathbf{x} \setminus [\mathbf{x}_k] = [\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \mathbf{x}_{k+1}, \dots, \mathbf{x}_N]$.

In a similar manner to the previous section, we denote by A_T the number of times an individual is duplicated by the size constraining mechanism before time T, and by B_T the number of times an individual is removed from the population by the size constraining mechanism. We also set $(\mathbf{X}(t))_{t\geq 0}$ to be the population process, $(|\mathbf{X}(t)|)_{t\geq 0}$ its size, and $(X_t)_{t\in\geq 0}$ a multi-dimensional birth and death process with (state dependent) transition rates β and δ. The notation $\mathbb{P}_{\mathbf{x}}$ (respectively $\mathbf{P}_x$) is used for the law of the population process $\mathbf{X}$ starting from $\mathbf{x} \in F$ (respectively for the law of X starting from $x \in E$). We denote as usual by $\mathbb{E}_x$ and $\mathbf{E}_x$ the associated expectations.

The first difficulty is to ensure that the constrained branching model defined above does not explode in finite time, where explosion is interpreted in the sense that infinitely many jumps of the system happen in finite time. Indeed, since κ is not assumed to be bounded, the rate at which new individuals are added to the system is not bounded. The following result provides a sufficient criterion for the non-explosion of the process but before stating it, let us first introduce some definitions that will be used in what follows.

We say that two functions V and κ are *co-monotone* if $(\kappa(x)-\kappa(y))(V(x)-V(y)) \geq 0$ for all $x, y \in E$. We say that V and κ are *almost co-monotone* if there exists a function κ' co-monotone with V such that $\kappa - \kappa'$ is bounded. We say that V *goes to infinity at infinity* if the set of $x \in E$ such that $V(x) < c$ is finite for all $c > 0$.

Proposition 13. *Assume that there exists a function $V : E \to [1, +\infty)$ almost co-monotone with κ, such that V tends to infinity at infinity, and such that, for all $x \in E$,*

$$\sum_{i=1}^{d} \left[\beta_i(x)(V(x+e_i) - V(x)) + \delta_i(x)(V(x-e_i) - V(x))\right] \leq CV(x), \tag{9}$$

for some constant $C \in \mathbb{R}$. Then the process with transition rates given by $\bar{L}$ is non-explosive.

Remark 14. The proof of this non-explosion result makes use of the extended infinitesimal generator of the particle system and of Dynkin's formula. As a consequence, its proof can be extended without too much effort to more general continuous time Markov processes, by replacing the left hand term in (9) by the extended infinitesimal generator of the more general process. One difficulty can appear though when the process is not a simple pure jump process on a discrete state space: to prove in a general context that $\bar{L}$ (adapted to fit the general individual dynamics of the particles) is indeed the extended infinitesimal generator of the particle system can be challenging.

Remark 15. More complex resampling/selection mechanisms could be considered with only slight modifications of the proof of Proposition 13. For instance, the main line of the proof applies as well if $N_{max} = +\infty$, or if we assume that, immediately after each branching time and with probability $1 + \frac{1}{N_{max}+1} - \frac{1}{N}$, a uniformly chosen individual is removed from the population, and/or that, immediately after each killing time and with probability $\frac{N_{min}-1}{N}$, a uniformly chosen individual is duplicated (where N is the population size just after the event). Note that these choices of parameters may be used to model an increasing competition between individuals when the size of the population increases.

Proof of Proposition 13. We define

$$\begin{aligned}\bar{V} :& F \to [1, +\infty)\\ :& \mathbf{x} = [\mathbf{x}_1, \ldots, \mathbf{x}_N] \mapsto \sum_{k=1}^{N} V(\mathbf{x}_k),\end{aligned}$$

so that $\bar{V}$ goes to infinity at infinity and for $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in F$,

$$\begin{aligned}
\bar{L}\bar{V}(\mathbf{x}) = \sum_{k=1}^{N}\sum_{i=1}^{d} &[\beta_i(\mathbf{x}_k)(V(\mathbf{x}_k + e_i) - V(\mathbf{x}_k)) + \delta_i(\mathbf{x}_k)(V(\mathbf{x}_k - e_i) - V(\mathbf{x}_k))] \\
&+ \sum_{k=1}^{N} (b(\mathbf{x}_k) - \kappa(\mathbf{x}_k))\, V(\mathbf{x}_k) \\
&+ \mathbf{1}_{N=N_{min}} \sum_{k=1}^{N} \kappa(\mathbf{x}_k) \frac{1}{N-1} \sum_{\ell=1, \ell\neq k}^{N} V(\mathbf{x}_l) \\
&- \mathbf{1}_{N=N_{max}} \sum_{k=1}^{N} b(\mathbf{x}_k) \frac{1}{N+1} \sum_{\ell=1}^{N} (1 + \mathbf{1}_{\ell=k}) V(\mathbf{x}_\ell).
\end{aligned}$$

Now let κ' be co-monotone with V such that $\kappa'' = \kappa - \kappa'$ is bounded. Then, using the fact that b is bounded, we deduce that

$$\begin{aligned}
\bar{L}\bar{V}(\mathbf{x}) &\leq (C + \|b\|_\infty)\overline{V}(\mathbf{x}) + \mathbf{1}_{N=N_{min}} \left(\frac{1}{N-1} \sum_{k=1}^{N} \kappa(\mathbf{x}_k) \sum_{\substack{\ell=1 \\ \ell\neq k}}^{N} V(\mathbf{x}_\ell) - \sum_{k=1}^{N} \kappa(\mathbf{x}_k) V(\mathbf{x}_k) \right) \\
&= (C + \|b\|_\infty)\overline{V}(\mathbf{x}) + \frac{\mathbf{1}_{N=N_{min}}}{N-1} \left(\sum_{k=1}^{N} \kappa(\mathbf{x}_k) \sum_{\ell=1}^{N} V(\mathbf{x}_\ell) - N \sum_{k=1}^{N} \kappa(\mathbf{x}_k) V(\mathbf{x}_k) \right) \\
&= (C + \|b\|_\infty)\overline{V}(\mathbf{x}) + \frac{N^2 \mathbf{1}_{N=N_{min}}}{N-1} \left[\left(\sum_{k=1}^{N} \frac{\kappa(\mathbf{x}_k)}{N} \right) \left(\sum_{k=1}^{N} \frac{V(\mathbf{x}_k)}{N} \right) - \sum_{k=1}^{N} \frac{\kappa(\mathbf{x}_k) V(\mathbf{x}_k)}{N} \right] \\
&\leq (C + \|b\|_\infty + 4\|\kappa''\|_\infty)\overline{V}(\mathbf{x}) \\
&\qquad + \frac{N^2 \mathbf{1}_{N=N_{min}}}{N-1} \left[\left(\sum_{k=1}^{N} \frac{\kappa'(\mathbf{x}_k)}{N} \right) \left(\sum_{k=1}^{N} \frac{V(\mathbf{x}_k)}{N} \right) - \sum_{k=1}^{N} \frac{\kappa'(\mathbf{x}_k) V(\mathbf{x}_k)}{N} \right] \\
&\leq (C + \|b\|_\infty + 4\|\kappa''\|_\infty)\overline{V}(\mathbf{x}), \qquad (10)
\end{aligned}$$

where we used the fact that κ' and V are co-monotone and the FKG inequality for the last inequality.

For all $n \geq 1$, let $U_n = \{\mathbf{x} \in F,\ \max_{k\in\{1,\dots,|\mathbf{x}|\}} \mathbf{x}_k > n\}$ and define the random variable $\theta_n := \inf\{t \geq 0,\ (X_t^1, \dots, X_t^N) \in U_n\}$. Note that θ_n is a stopping time since the constrained branching process is càdlàg. Moreover, denoting by C_n the set of points in F that can be reached in one jump from $F \setminus U_n$, we observe that C_n is bounded and hence that $\bar{V}(X_{t\wedge\theta_n})$ is almost surely uniformly bounded in $t \geq 0$. In particular, we deduce, using Dynkin's formula and the inequality (10), that, for all $n \geq 0$, all $t \geq 0$ and all $\mathbf{x} \in F$,

$$\mathbb{E}_{\mathbf{x}}(\bar{V}(\mathbf{X}_{t\wedge\theta_n})) \leq \bar{V}(\mathbf{x}) + (C + \|b\|_\infty + 4\|\kappa''\|_\infty) \int_0^t \mathbb{E}_x(\bar{V}(\mathbf{X}_{s\wedge\theta_n}))\, \mathrm{d}s.$$

Hence, using Grönwall's inequality, we deduce that

$$\mathbb{E}_{\mathbf{x}}(\bar{V}(\mathbf{X}_{t\wedge\theta_n})) \leq \bar{V}(\mathbf{x})e^{t(C+\|b\|_\infty+4\|\kappa''\|_\infty)}.$$

Since $\{\mathbf{X}_{t\wedge\theta_n} \in U_n, \theta_n \leq t\} = \{\theta_n \leq t\}$ up to a negligible event, we conclude that

$$\inf_{\mathbf{y}\in U_n} \bar{V}(\mathbf{y})\mathbb{P}_{\mathbf{x}}(\theta_n \leq t) \leq \bar{V}(\mathbf{x})\, e^{t(C+\|b\|_\infty+4\|\kappa''\|_\infty)}.$$

Denote by τ_∞ the explosion time of the process. Since the jump rate of the process is bounded up to time θ_n, we deduce that $\theta_n \leq \tau_\infty$ almost surely. Hence

$$\mathbb{P}_{\mathbf{x}}(\tau_\infty \leq t) \leq \frac{\bar{V}(\mathbf{x})e^{t(C+\|b\|_\infty+4\|\kappa''\|_\infty)}}{\inf_{\mathbf{y}\in U_n}\bar{V}(\mathbf{y})} \xrightarrow[n\to+\infty]{} 0,$$

so that $\tau_\infty = +\infty$ almost surely. This concludes the proof. □

Since the process defined in this section is a particular instance of the process described in Section 2.1, the following corollary is a direct consequence of Proposition 13 and Theorem 6.

Corollary 16. *Under the assumptions of Proposition 13, we have, for all $T > 0$, all bounded measurable functions $f : \mathbb{Z}_+^d \to \mathbb{R}$ and all initial position $\mathbf{X}(0) \in F$,*

$$\mathbb{E}_{\mathbf{X}(0)}\left[\Pi_T^A\Pi_T^B \sum_{k=1}^{|\mathbf{X}(t)|} f(\mathbf{X}(t)_k)\right] = \sum_{k=1}^{|\mathbf{X}(0)|} \mathbf{E}_{\mathbf{X}(0)_k}\left[f(X_t)\exp\left(\int_0^t (b(X_s)-\kappa(X_s))\,\mathrm{d}s\right)\right]$$

where

$$\Pi_T^A = \left(\frac{N_{min}-1}{N_{min}}\right)^{A_T} \text{ and } \Pi_T^B = \left(\frac{N_{max}+1}{N_{max}}\right)^{B_T}.$$

In addition,

$$\left\|\frac{\sum_{k=1}^{|\mathbf{X}(0)|}\mathbf{E}_{\mathbf{X}(0)_k}\left[f(X_t)\exp\left(\int_0^t(b(X_s)-\kappa(X_s))\,\mathrm{d}s\right)\right]}{\sum_{k=1}^{|\mathbf{X}(0)|}\mathbf{E}_{\mathbf{X}(0)_k}\left[\exp\left(\int_0^t(b(X_s)-\kappa(X_s))\,\mathrm{d}s\right)\right]} - \frac{\sum_{k=1}^{|\mathbf{X}(t)|} f(\mathbf{X}(t)_k)}{|\mathbf{X}(t)|}\right\|_2$$
$$\leq C_T \frac{\|f\|_\infty}{\frac{1}{|\mathbf{X}(0)|}\sum_{k=1}^{|\mathbf{X}(0)|}\mathbf{E}_{\mathbf{X}(0)_k}\left[\exp\left(\int_0^t(b(X_s)-\kappa(X_s))\,\mathrm{d}s\right)\right]}\frac{1}{\sqrt{|\mathbf{X}(0)|}}.$$

Our aim is now to study the long-time behaviour of the (normalised) empirical distribution of the health parameter across the population. We assume that there exist constants $C_1 > 0, C_2 > 0, \eta_0 > 0, n_0 \geq 0$ such that, for all $x \in \mathbb{Z}_+^d$ satisfying $|x| \geq n_0$,

$$\sum_{i=1}^d (\delta_i(x) - \beta_i(x)) \geq C_1\,|x|^{1+\eta_0} \tag{11}$$

$$\sum_{i=1}^d (\delta_i(x) - \beta_i(x)) \geq C_2(\|b\|_\infty - b(x) + \kappa(x))|x|. \tag{12}$$

Proposition 17. *Assume that* (11) *and* (12) *hold true and that* X *is irreducible. Then there exists a probability distribution* ν_X *on* $\mathbb{Z}_+^d$ *such that, for all probability measures* μ *on* $\mathbb{Z}_+^d$ *and all bounded measurable functions* f *on* $\mathbb{Z}_+^d$,

$$\left|\frac{\mathbf{E}_\mu\left(f(X_t)\exp\left(\int_0^t b(X_s)-\kappa(X_s)\,\mathrm{d}s\right)\right)}{\mathbf{E}_\mu\left(\exp\left(\int_0^t b(X_s)-\kappa(X_s)\,\mathrm{d}s\right)\right)}-\nu_X(f)\right|\leq C_X\,e^{-\gamma_X t}\,\|f\|_\infty, \tag{13}$$

for some constants $C_X\geq 1$ *and* $\gamma_X>0$. *In addition, there exists a bounded function* $\eta_X:\mathbb{Z}_+^d\to(0,+\infty)$ *and* $\lambda_X>-\|b\|_\infty$ *such that for all probability measures* μ *on* $\mathbb{Z}_+^d$,

$$\mathbf{E}_\mu\left(\eta_X(X_t)\exp\left(\int_0^t b(X_s)-\kappa(X_s)\,\mathrm{d}s\right)\right)=e^{-\lambda_X t}\mu(\eta_X),\ \forall t\geq 0.$$

Proof of Proposition 17. First observe that, for all bounded measurable functions $f:\mathbb{Z}_+^d\to\mathbb{R}$,

$$\mathbf{E}_\mu\left(f(X_t)\exp\left(\int_0^t b(X_s)-\kappa(X_s)\,\mathrm{d}s\right)\right)=e^{\|b\|_\infty t}\,\mathbf{E}^b_\mu\left(f(X^b_t)\mathbf{1}_{t<\tau^b_\partial}\right),$$

where $\mathbf{P}^b_\mu$ and $\mathbf{E}^b_\mu$ denote the law and associated expectation of the birth and death process X^b with birth and death parameters β and δ, killed at rate $\kappa^b:=\|b\|_\infty-b+\kappa$ at a time denoted by τ^b_∂. We now prove that there exist ν_X, $C_X>0$ and $\gamma_X>0$ such that, for all bounded measurable functions $f:\mathbb{Z}_+^d\to\mathbb{R}$

$$\left|\mathbf{E}^b_\mu\left(f(X^b_t)\mid t<\tau^b_\partial\right)-\nu_X(f)\right|\leq C_x\,e^{-\gamma_X t}\|f\|_\infty, \tag{14}$$

which is equivalent to (13).

In order to prove (14), we check that the assumptions of [11, Theorem 2.4] hold true. Since the proof is similar to Theorem 4.1 in the above reference, we only recall the main ingredients. Set $W(x)=|x|$ for all $x\in\mathbb{Z}_+^d$. Assumption (11) implies that, for all $x\in\mathbb{Z}_+^p$ large enough,

$$\sum_{i=1}^d\left[\beta_i(x)(W(x+e_i)-W(x))+\delta_i(x)(W(x-e_i)-W(x))\right]\leq -C_1W(x)^{1+\eta_0}.$$

This classically entails that the birth and death process on $\mathbb{Z}_+^d$ with birth rates β and death rates δ (without absorption) is non-explosive and comes down from infinity. This implies that the process is regularly absorbed (as defined in (2.1), p.55 of [11]) and that Assumption 3 in [11] is satisfied.

Moreover, choosing appropriately $a_1>1$, $a_2>1$ and $\varepsilon>0$, and setting

$$V(x)=\sum_{k=1}^{|x|}\frac{1}{k^{a_1}}\text{ and }\varphi(x)=\sum_{k=|x|+1}^{+\infty}\frac{1}{k^{a_2}},$$

one deduces, using the same computations as in Section 4 in [11], that Assumption 1 of [11] is satisfied. The only difference is that the infinitesimal generator that one needs to consider in the present case is

$$Lf(x)=\sum_{i=1}^d\left[\beta_i(x)(f(x+e_i)-f(x))+\delta_i(x)(f(x-e_i)-f(x))\right]-\kappa^b(x)f(x).$$

Finally, Assumption 2 in the above reference is immediate since the process is irreducible in a locally finite state space. This shows that [11, Theorem 2.4] applies to X^b, which concludes the proof of (13).

The last assertion of the proposition follows from [10, Proposition 2.3]: it implies that there exists $\lambda_0 > 0$ and a bounded positive function η_X such that

$$\mathbf{E}^b_\mu(\eta_X(X^b_t)) = e^{-\lambda_0 t}\eta_X.$$

Multiplying by $e^{\|b\|_\infty t}$ and setting $\lambda_X = \lambda_0 - \|b\|_\infty$ concludes the proof. □

We consider now the situation where κ is bounded and where N_{min} is large, and prove that the normalised empirical distribution $\dfrac{m_T}{m_T(\mathbf{1}_E)}$ of the population health's parameter converges toward ν_X when N_{min}, N_{max} and T go to infinity.

Proposition 18. *Assume that the conditions of Proposition 17 hold and that κ is bounded. Then, denoting by m_0 the initial empirical measure of the population's health parameter and setting $\widetilde{m}_0 = m_0/m_0(\mathbf{1}_E)$, we have*

$$\sup_{T\in[0,+\infty)} \left\| \frac{m_T(f)}{m_T(\mathbf{1}_E)} - \frac{\mathbf{E}_{\widetilde{m}_0}\left[f(X_t)\exp\left(\int_0^t (b(X_s)-\kappa(X_s))\,\mathrm{d}s\right)\right]}{\mathbf{E}_{\widetilde{m}_0}\left[\exp\left(\int_0^t (b(X_s)-\kappa(X_s))\,\mathrm{d}s\right)\right]} \right\|_2 \leq (2C_X+C)N_{min}^{-\alpha/2}\,\|f\|_\infty,$$

where $\alpha = \gamma_X/((c+1)\|b\|_\infty + \|\kappa\|_\infty + \gamma_X) \in (0,1)$, C_X, γ_X are from Proposition 17, and C, c are from Theorem 6. In particular,

$$\left\| \frac{m_T(f)}{m_T(\mathbf{1}_E)} - \nu_X(f) \right\|_2 \leq \left[(2C_X + C)N_{min}^{-\alpha/2}\,\|f\|_\infty + C_X\, e^{-\gamma_X T}\right]\,\|f\|_\infty.$$

Remark 19. This result can be adapted to include the case where X is a Markov process with *hard* killing at a boundary when the property obtained in Proposition 17 is known to hold true (see for instance [23, 10, 9, 36, 25]), provided one can prove that the particle system does not degenerate toward the boundary. However this can be challenging, see for instance [33, 56, 5].

Proof of Proposition 18. We assume without loss of generality that $\|f\|_\infty = 1$ and that $N_{min} > \sqrt{N_{max}}$ (otherwise the result is trivial). We define the semigroup Q by

$$Q_t f(x) = \mathbf{E}_x\left[f(X_t)\exp\left(\int_0^t (b(X_s)-\kappa(X_s))\,\mathrm{d}s\right)\right].$$

We have, by Proposition 17 applied first to $\mu = m_0 Q_{T-t}/m_0 Q_{T-t}\mathbf{1}_E$ and then to $\widetilde{m}_{T-t} = m_{T-t}/m_{T-t}\mathbf{1}_E$, and by Theorem 6, for all $0 \leq t \leq T$,

$$\left\| \frac{m_0 Q_T f}{m_0 Q_T \mathbf{1}_E} - \frac{m_T(f)}{m_T(\mathbf{1}_E)} \right\|_2 \leq \left\| \frac{m_0 Q_{T-t} Q_t f}{m_0 Q_{T-t} Q_t \mathbf{1}_E} - \frac{m_{T-t}(Q_t f)}{m_{T-t}(Q_t \mathbf{1}_E)} \right\|_2 + \left\| \frac{m_{T-t}(Q_t f)}{m_{T-t}(Q_t \mathbf{1}_E)} - \frac{m_T(f)}{m_T(\mathbf{1}_E)} \right\|_2$$

$$\leq 2C_X e^{-\gamma_X t} + C\exp(c\|b\|_\infty t)\left\| \frac{\|Q_t\mathbf{1}_E\|_\infty}{\widetilde{m}_{T-t}Q_t\mathbf{1}_E}\frac{1}{\sqrt{m_{T-t}\mathbf{1}_E}} \right\|_2, \tag{15}$$

where

$$\left\| \frac{\|Q_t \mathbf{1}_E\|_\infty}{\widetilde{m}_{T-t} Q_t \mathbf{1}_E} \frac{1}{\sqrt{m_{T-t}\mathbf{1}_E}} \right\|_2 \leq \frac{\exp(\|b\|_\infty t)}{\exp(-\|\kappa\|_\infty t)} \frac{1}{\sqrt{N_{min}}}.$$

Set

$$t_0 = \frac{\ln\left(\sqrt{N_{min}}\right)}{(c+1)\|b\|_\infty + \|\kappa\|_\infty + \gamma_X},$$

so that $e^{-\gamma_X t_0} = e^{((c+1)\|b\|_\infty + \|\kappa\|_\infty)t_0} / \sqrt{N_{min}}$. If $T \geq t_0$, then choosing $t = t_0$ in (15), we get

$$\begin{aligned} \left\| \frac{m_0 Q_T f}{m_0 Q_T \mathbf{1}_E} - \frac{m_T(f)}{m_T(\mathbf{1}_E)} \right\|_2 &\leq (2C_X + C) e^{-\gamma_X t_0} \\ &= (2C_X + C) N_{min}^{-\alpha/2}, \end{aligned}$$

with $\alpha = \dfrac{\gamma_X}{(c+1)\|b\|_\infty + \|\kappa\|_\infty + \gamma_X}$. For $T \leq t_0$, Theorem 6 implies that

$$\begin{aligned} \left\| \frac{m_0 Q_T f}{m_0 Q_T \mathbf{1}_E} - \frac{m_T(f)}{m_T(\mathbf{1}_E)} \right\|_2 &\leq C \exp\left((c+1)\|b\|_\infty t_0 + \|\kappa\|_\infty t_0\right) \frac{1}{\sqrt{N_{min}}} \\ &= C N_{min}^{-\alpha/2} \end{aligned}$$

which concludes the proof of the first inequality in Proposition 18. The second inequality is then a straightforward consequence of the first inequality and of Proposition 17. □

3.2 Piecewise deterministic Markov processes

Due to Assumption 1 in Section 2, the results presented thus far do not hold for the case where $(X_t)_{t\geq 0}$ is a piecewise deterministic Markov process (PDMP) and E is a bounded domain with absorbing boundary conditions, since two independent copies of the process may hit ∂E at the same time. In this section, we consider a particular example, namely the neutron transport equation (NTE), in order to discuss a way to avoid this problem.

The NTE is a balance equation that describes the behaviour of neutrons in a fissile medium such as a nuclear reactor. In such systems neutrons move in straight lines with a fixed speed until they either come into contact with the boundary of the reactor, at which point they are absorbed, or they collide with the nucleus of an atom. When the latter occurs, either the neutron undergoes a scattering event where the particle bounces off the nucleus and continues its motion but with a new velocity, or a fission event occurs where the collision causes new neutrons to be produced with identical spatial positions, but potentially different velocities.

In this setting, we have $E = D \times V$ where $D \subset \mathbb{R}^3$ is the spatial domain, which we assume to be open and bounded with smooth boundary, ∂D, and $V = \{\upsilon \in \mathbb{R}^3 : \upsilon_{min} < |\upsilon| < \upsilon_{max}\}$ is the velocity domain, where $0 < \upsilon_{min} \leq \upsilon_{max} < \infty$. The NTE can then be stated as follows:

$$\begin{aligned} \frac{\partial}{\partial t}\psi_t(r,\upsilon) &= \upsilon \cdot \nabla \psi_t(r,\upsilon) - (\sigma_{\mathtt{s}}(r,\upsilon) + \sigma_{\mathtt{f}}(r,\upsilon))\psi_t(r,\upsilon) \\ &\quad + \sigma_{\mathtt{s}}(r,\upsilon)\int_V \psi_t(r,\upsilon')\pi_{\mathtt{s}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon' + \sigma_{\mathtt{f}}(r,\upsilon)\int_V \psi_t(r,\upsilon')\pi_{\mathtt{f}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon', \end{aligned} \quad (16)$$

where

$\sigma_{\mathtt{s}}(r,\upsilon)$: the rate at which scattering occurs from incoming velocity υ at position r,

$\sigma_{\mathtt{f}}(r,\upsilon)$: the rate at which fission occurs from incoming velocity υ at position r,

$\pi_{\mathtt{s}}(r,\upsilon,\upsilon')$: probability density that an incoming velocity υ at position r scatters to an outgoing velocity, with probability υ' satisfying $\int_V \pi_{\mathtt{s}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon' = 1$, and

$\pi_{\mathtt{f}}(r,\upsilon,\upsilon')$: density of expected neutron yield at velocity υ' from fission with incoming velocity υ satisfying $\int_V \pi_{\mathtt{f}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon' < \infty$.

Moreover, we impose the following initial and boundary conditions

$$\left.\begin{aligned} &\psi_0(r,\upsilon) = g(r,\upsilon) && \text{for } r \in D, \upsilon \in V,\\ &\psi_t(r,\upsilon) = 0 && \text{for } r \in \partial D \text{ if } \upsilon\cdot\mathbf{n}_r > 0, \end{aligned}\right\} \tag{17}$$

where $\mathbf{n}_r$ is the outward facing normal of D at $r \in \partial D$ and $g : D \times V \to [0,\infty)$ is a bounded, measurable function.

Under the assumptions,

(A1) $\sigma_{\mathtt{s}}$, $\sigma_{\mathtt{f}}$, $\pi_{\mathtt{s}}$ and $\pi_{\mathtt{f}}$ are uniformly bounded away from infinity.

(A2) $\inf_{r\in D,\upsilon,\upsilon'\in V}(\sigma_{\mathtt{s}}(r,\upsilon)\pi_{\mathtt{s}}(r,\upsilon,\upsilon') + \sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon')) > 0$,

it was shown in [14, 36] that the NTE can be modelled using a weighted PDMP, also known as the neutron random walk (NRW). More precisely, setting

$$\alpha(r,\upsilon) = \sigma_{\mathtt{s}}(r,\upsilon) + \sigma_{\mathtt{f}}(r,\upsilon)\int_V \pi_{\mathtt{f}}(r,\upsilon,\upsilon')\mathrm{d}\upsilon' \tag{18}$$

$$\pi(r,\upsilon,\upsilon') = (\alpha(r,\upsilon))^{-1}\left[\sigma_{\mathtt{s}}(r,\upsilon)\pi_{\mathtt{s}}(r,\upsilon,\upsilon') + \sigma_{\mathtt{f}}(r,\upsilon)\pi_{\mathtt{f}}(r,\upsilon,\upsilon')\right], \tag{19}$$

$$\beta(r,\upsilon) = \alpha(r,\upsilon) - \sigma_{\mathtt{s}}(r,\upsilon) - \sigma_{\mathtt{f}}(r,\upsilon), \tag{20}$$

we define the NRW $((R_t,\Upsilon_t)_{t\geq 0}, \mathbf{P}_{(r,\upsilon)})$ as follows. From an initial configuration $(r,\upsilon) \in D\times V$, the particle will propagate linearly until it is either absorbed at the boundary of the domain or, at rate α the process scatters off a nucleus and a new velocity is chosen according to π. Then, for a bounded, measurable function g, and $(r,\upsilon) \in D \times V$

$$\psi_t[g](r,\upsilon) := \mathbf{E}_{(r,\upsilon)}\left[\mathrm{e}^{\int_0^t \beta(R_s,\Upsilon_s)\mathrm{d}s} g(R_t,\Upsilon_t)\mathbf{1}_{(t<\tau_D)}\right],$$

solves (16)–(17), where τ_D is the first exit time of the NRW from D.

It was also shown in [36] that if (A1) and (A2) are satisfied, and that if there exists an $\varepsilon > 0$ such that

(B1) $D_\varepsilon := \{r \in D : \inf_{y\in\partial D}|r-y| > \varepsilon \mathtt{v}_{\mathtt{max}}\}$ is non-empty and connected,

(B2) there exist $0 < s_\varepsilon < t_\varepsilon$ and $\gamma > 0$ such that, for all $r \in D\backslash D_\varepsilon$, there exists $K_r \subset V$ measurable such that $\mathrm{Vol}(K_r) \geq \gamma > 0$ and for all $\upsilon \in K_r$, $r + \upsilon s \in D_\varepsilon$ for every $s \in [s_\varepsilon, t_\varepsilon]$ and $r + \upsilon s \notin \partial D$ for all $s \in [0, s_\varepsilon]$,

the semigroup $(\psi_t)_{t\geq 0}$ exhibits the following Perron Frobenius behaviour.

Theorem ([36]). *There exists a $\lambda_* \in \mathbb{R}$, a positive right eigenfunction $\varphi \in L_\infty^+(D \times V)$ and a left eigenmeasure which is absolutely continuous with respect to Lebesgue measure on $D \times V$ with density $\tilde{\varphi} \in L_\infty^+(D \times V)$, both having associated eigenvalue $\mathrm{e}^{\lambda_* t}$, and such that φ (resp. $\tilde{\varphi}$) is uniformly (resp. a.e. uniformly) bounded away from zero on each compactly embedded subset of $D \times V$. In particular, for all $g \in L_\infty^+(D \times V)$,*

$$\langle \tilde{\varphi}, \psi_t[g] \rangle = \mathrm{e}^{\lambda_* t} \langle \tilde{\varphi}, g \rangle \quad (\text{resp. } \psi_t[\varphi] = \mathrm{e}^{\lambda_* t} \varphi), \quad t \geq 0. \tag{21}$$

Moreover, there exists $\varepsilon > 0$ such that

$$\sup_{g \in L_\infty^+(D \times V): \|g\|_\infty \leq 1} \left\| \mathrm{e}^{-\lambda_* t} \varphi^{-1} \psi_t[g] - \langle \tilde{\varphi}, g \rangle \right\|_\infty = O(\mathrm{e}^{-\varepsilon t}), \qquad t \geq 0. \tag{22}$$

This theorem therefore provides a way to build Monte Carlo simulations to estimate the eigenvalue λ_* and the eigenfunctions φ and $\tilde{\varphi}$ via the NRW. We refer the reader to [15] for further details. However, although simulating a single weighted path has advantages over simulating an entire tree of neutrons, the transience of this process means that many of the particles exit the domain relatively quickly and therefore, a large number of simulations are required in order to obtain information about the system. This also leads to problems with the variance of the estimators. In order to deal with this problem, the notion of 'h-transform' was developed in [15], where one biases the NRW to prevent it from exiting the domain.

More precisely, let h be a bounded positive function on $D \times V$ such that $h = 0$ on ∂D and define the following change of measure,

$$\left. \frac{\mathrm{d}\mathbf{P}^h_{(r,\upsilon)}}{\mathrm{d}\mathbf{P}_{(r,\upsilon)}} \right|_{\sigma((R_s, \Upsilon_s), s \leq t)} := \exp\left(-\int_0^t \frac{\mathtt{J}h(R_s, \Upsilon_s)}{h(R_s, \Upsilon_s)} \mathrm{d}s \right) \prod_{i=1}^{N_t} \frac{h(R_{T_i}, \Upsilon_{T_i})}{h(R_{T_i}, \Upsilon_{T_{i-1}})} \mathbf{1}_{(t < \tau^D)}, \tag{23}$$

where

$$\mathtt{J}g(r, \upsilon) = \alpha(r, \upsilon) \int_V [g(r, \upsilon') - g(r, \upsilon)] \pi(r, \upsilon, \upsilon') \mathrm{d}\upsilon', \tag{24}$$

$(T_i)_{i \geq 0}$ are the scatter times of the NRW with $T_0 = 0$, and $N_t = \sup\{i : T_i \leq t\}$.

Then, $((R, \Upsilon), \mathbf{P}^h)$ also defines a NRW but with scattering operator

$$\mathtt{J}_h g(r, \upsilon) = \alpha(r, \upsilon) \int_V [g(r, \upsilon') - g(r, \upsilon)] \frac{h(r, \upsilon')}{h(r, \upsilon)} \pi(r, \upsilon, \upsilon') \mathrm{d}\upsilon'. \tag{25}$$

The idea is that the factor $h(r, \upsilon')/h(r, \upsilon)$ in the above integral forces the NRW to scatter when it is approaching the boundary, thus preventing it from being killed. Thus, if we can show that this process satisfies Assumption 1, we may use the BBMMI process in conjunction with the h-transformed process to estimate the leading eigentriple $(\lambda_*, \varphi, \tilde{\varphi})$, for example. Indeed, it is straightforward to show that the leading eigentriple of the h-transformed process is given by $(\lambda_*, \varphi/h, h\tilde{\varphi})$ and so we may use the BBMMI process to estimate this triple, and hence obtain an estimate for the original quantities $(\lambda_*, \varphi, \tilde{\varphi})$. The rest of this section is dedicated to verifying Assumption 1 for the transformed process. We refer the reader to Section 3.3 for the numerical aspects.

To this end, fix $\delta = 1/(2\|\alpha\|_\infty)$ and let $\phi : [0, +\infty) \to [0, \delta]$ be a regular non-decreasing function such that $\phi(x) = x$ for all $x \in [0, \delta/2]$ and $\phi(x) = \delta$ for all $x \geq \delta$. We set

$$h(r, \upsilon) = \phi(\kappa_{r,\upsilon}).$$

Using the following result, one can express expectations with respect to $\mathbf{P}$ as Feynman-Kac formulas related to $\mathbf{P}^h$. In what follows, we say that (r, υ) is in a vicinity of ∂D if $h(r, \upsilon) < \delta/2$.

Lemma 20. *Suppose (A1), (A2) hold. A process with law $\mathbf{P}^h_{(r,\upsilon)}$ does not hit the boundary with probability one. Moreover, setting $b = \left(\frac{Lh}{h}\right)_+$ and $\kappa = \left(\frac{Lh}{h}\right)_-$, the function b is bounded on $D \times V$ and κ is locally bounded. Let $\mathbf{P}^\kappa_{r,\upsilon}$ be the law of a process with law $\mathbf{P}^h_{r,\upsilon}$ with additional soft killing at rate κ, then we have*

$$\mathbf{E}_{(r,\upsilon)}\left[f(R_t,\Upsilon_t)\,\mathbf{1}_{(t<\tau^D)}\right] = h(r,\upsilon)\,\mathbf{E}^\kappa_{(r,\upsilon)}\left[\exp\left(\int_0^t b(R_s,\Upsilon_s)\,\mathrm{d}s\right)\frac{f(R_t,\Upsilon_t)}{h(R_t,\Upsilon_t)}\mathbf{1}_{(t<\tau_\kappa)}\right], \tag{26}$$

where τ_κ denotes the soft killing time of the process with law $\mathbf{P}^\kappa_{r,\upsilon}$.

Proof. We have

$$\begin{aligned}&\psi_t[g](r,\upsilon)\\ &= \mathbf{E}^h_{(r,\upsilon)}\left[\exp\left(\int_0^t \frac{\mathtt{J}h(R_s,\Upsilon_s)}{h(R_s,\Upsilon_s)} + \beta(R_s,\Upsilon_s)\mathrm{d}s\right)\prod_{i=1}^{N_t}\frac{h(R_{T_i},\Upsilon_{T_{i-1}})}{h(R_{T_i},\Upsilon_{T_i})}g(R_t,\Upsilon_t)\mathbf{1}_{(t<\tau^D)}\right]\end{aligned} \tag{27}$$

Note that

$$\inf_{r\in D,\upsilon\in V}\lim_{s\to\kappa^D_{r,\upsilon}}\frac{|\upsilon|(\kappa^D_{r,\upsilon}-s)}{h(r+\upsilon s,\upsilon)} > 0, \tag{28}$$

where $\kappa^D_{r,\upsilon} := \inf\{t>0 : r+\upsilon t \notin D\}$. Hence, Theorem 7.1 in [15] implies that, under $\mathbf{P}^h_{(r,\upsilon)}$, $r\in D, \upsilon\in V$, (R,Υ) does not hit the boundary with probability one.

Recall that the infinitesimal generator associated to $((R_t,\Upsilon_t)_{t\geq 0},\mathbf{P})$ is given by

$$Lf(r,\upsilon) = \mathtt{T}f + \alpha(r,\upsilon)\int_V (f(r,\upsilon') - f(r,\upsilon))\,\pi(r,\upsilon,\upsilon'),\mathrm{d}\upsilon,$$

where $\mathtt{T}$ is defined for regular functions by $\mathtt{T}f(r,\upsilon) = \upsilon\cdot\nabla_r f(r,\upsilon)$ and, more generally and when the limit exists, by

$$\mathtt{T}f(r,\upsilon) = \lim_{\varepsilon\to 0^+}\frac{f(r+\varepsilon\upsilon) - f(r,\upsilon)}{\varepsilon}.$$

In particular, the conditions of Remark 7.1 of [15] are satisfied and hence

$$\left.\frac{\mathrm{d}\mathbf{P}^h_{(r,\upsilon)}}{\mathrm{d}\mathbf{P}_{(r,\upsilon)}}\right|_{\sigma((R_s,\Upsilon_s),s\leq t)} = \exp\left(-\int_0^t \frac{Lh(R_s,\Upsilon_s)}{h(R_s,\Upsilon_s)}\mathrm{d}s\right)\frac{h(R_t,\Upsilon_t)}{h(r,\upsilon)}\,\mathbf{1}_{(t<\tau^D)}. \tag{29}$$

Finally, we obtain, for all measurable functions $f: D\times V \to [0,+\infty)$,

$$\mathbf{E}_{(r,\upsilon)}\left[f(R_t,\Upsilon_t)\,\mathbf{1}_{(t<\tau^D)}\right] = h(r,\upsilon)\,\mathbf{E}^h_{(r,\upsilon)}\left[\exp\left(\int_0^t \frac{Lh(R_s,\Upsilon_s)}{h(R_s,\Upsilon_s)}\mathrm{d}s\right)\frac{f(R_t,\Upsilon_t)}{h(R_t,\Upsilon_t)}\right] \tag{30}$$

where have used the fact that $\mathbf{1}_{(t<\tau^D)} = 1$, $\mathbf{P}^h_{(r,\upsilon)}$-almost surely, deduced from (28). Now, introducing $b = \left(\frac{Lh}{h}\right)_+$ and $\kappa = \left(\frac{Lh}{h}\right)_-$, and considering the law $\mathbf{P}^\kappa_{r,\upsilon}$ of a process with law $\mathbf{P}^h_{r,\upsilon}$ with soft killing at rate κ, we deduce that

$$\mathbf{E}_{(r,\upsilon)}\left[f(R_t,\Upsilon_t)\,\mathbf{1}_{(t<\tau^D)}\right] = h(r,\upsilon)\,\mathbf{E}^\kappa_{(r,\upsilon)}\left[\exp\left(\int_0^t b(R_s,\Upsilon_s)\,\mathrm{d}s\right)\frac{f(R_t,\Upsilon_t)}{h(R_t,\Upsilon_t)}\mathbf{1}_{(t<\tau_\kappa)}\right], \tag{31}$$

where τ_κ denotes the soft killing time of the process (with law $\mathbf{P}^h_{r,\upsilon}$ with soft killing at rate κ).

Now, we observe that, in a vicinity of ∂D,

$$\begin{aligned} Lh(r,\upsilon) &= -1 + \alpha(r,\upsilon)\int_V (h(r,\upsilon') - h(r,\upsilon))\,\pi(r,\upsilon,\upsilon')\,\mathrm{d}\upsilon' \\ &\le -1 + \|\alpha\|_\infty \delta = -1/2 \end{aligned} \tag{32}$$

and hence that

$$\sup_{(r,\upsilon)\in D\times V} \frac{Lh(r,\upsilon)}{h(r,\upsilon)} < +\infty.$$

This implies that b is bounded. The fact that κ is locally bounded is an immediate consequence of the regularity of h. □

Our aim is now to apply the results of Section 2.2 to the Feynman-Kac expression (26) in order to obtain information about $\mathbf{P}$. Assumption 1 is clearly satisfied by a process with law $\mathbf{P}^\kappa_{(r,\upsilon)}$, and it only remains to check that Assumption 2 holds true in order to apply Theorem 6. This is the purpose of the following result. In what follows, we refer to a bounded BBMMI as any BBMMI whose selection mechanism imposed by the parameters p^i results in a size constrained process, i.e. there is a constant N_{max} such that $\bar{N}_t \le N_{max}$ for all $t \ge 0$.

Proposition 21. *Under the assumptions (A1), (A2), (B1) and (B2), any bounded BBMMI driven by $\mathbf{P}^\kappa_{(r,\upsilon)}$ satisfies Assumption 2.*

Proof. We denote by $\sigma_1 < \sigma_2 < ...$ the sequence of times at which events occur for a given particle in the system (each time is either a scattering, a branching or a soft killing). Note that each time σ_n is well defined since, almost surely, the number of event occurring before a time T goes to infinity when $T \to +\infty$. Since one can use the strong Markov property at the birth time of the given particle, we assume without loss of generality that the particle is already alive at time 0.

From the expression of the scattering operator $\underline{\mathrm{J}}_h$ in (25), one observes that the scattering rate toward a direction in a set $V' \subset V$ is given by $\int_{V'} \alpha(r,\upsilon)\frac{h(r,\upsilon')}{h(r,\upsilon)}\pi(r,\upsilon,\upsilon')\mathrm{d}\upsilon'$. But, under the assumptions (B1–2), there exists $\varepsilon > 0$, $\beta > 0$ and $\eta > 0$ such that, for any point $r \in D$, there exists V_r with $r + [0,\beta]V_r \subset D$ and $\mathrm{Leb}_d((r+[0,\beta]V_r) \cap \{r' \in D,\ d(r',\partial D) \ge \varepsilon\}) > \eta$, where Leb_d denotes the d-dimensional Lebesgue measure. In particular, using the fact that $h(r,\upsilon')$ is lower bounded on V_r (uniformly over $r \in D$). We deduce that there exists a constant $\underline{s} > 0$ such that the scattering rate toward a direction in V_r is lower bounded by $\underline{s}/h(r,\upsilon)$, uniformly in $(r,\upsilon) \in D \times V$.

Using (32) to bound Lh from below, there exists a constant $\bar{c}$ such that the total rate at which a particle undergoes an event (either a scattering, a branching or a (soft) killing) is upper bounded by $\bar{c}/h(r,\upsilon)$, uniformly in $(r,\upsilon) \in D \times V$. Hence each event has a probability greater than $\underline{s}/\bar{c}$ to be a scattering event toward a direction in V_r, independently of the past of the process, and in particular independently of the time event σ_n: formally, for all $n \ge 0$,

$$\mathbf{P}^\kappa\left(\sigma_n \text{ is a scattering event and } (R_{\sigma_n}, \Upsilon_{\sigma_n}) \in \{R_{\sigma_n -}\} \times V_{R_{\sigma_n -}} \mid \sigma_n, R_{\sigma_n -}, \Upsilon_{\sigma_n -}\right) \ge \underline{s}/\bar{c} > 0,$$

where (R,Υ) denotes the position and direction of the given particle.

Now, using the fact that the killing rate and the scattering rate are uniformly bounded in $\{r' \in D,\ d(r',\partial D) \ge \varepsilon\}$ and the fact that the total branching rate of the system is uniformly bounded, we deduce that there exists a constant $\underline{p} > 0$ and a time $T > 0$ such that, for all $r \in D$ and $\upsilon \in V_r$, $\mathbb{P}(\sigma_1 \geq T) \geq \underline{p}$.

Hence, we deduce from the strong Markov property that

$$\begin{aligned}
\mathbf{P}^{\kappa}(\sigma_{n+1} < T) &\leq \mathbf{E}^{\kappa}\left[\mathbf{1}_{\sigma_n < T}\, \mathbf{P}^{\kappa}_{X_{\sigma_n}}(\sigma_1 < T)\right] \\
&\leq \mathbf{E}^{\kappa}\left[\mathbf{1}_{\sigma_n < T}\left(1 - \mathbf{1}_{\Upsilon_{\sigma_n} \in V_{R_{\sigma_n}}} \mathbf{P}^{\kappa}_{X_{\sigma_n}}(\sigma_1 < T)\right)\right] \\
&\leq \mathbf{E}^{\kappa}\left[\mathbf{1}_{\sigma_n < T}\left(1 - \mathbf{1}_{\Upsilon_{\sigma_n} \in V_{R_{\sigma_n}}} \underline{p}\right)\right] \\
&\leq \mathbf{P}^{\kappa}(\sigma_n < T)(1 - \underline{s}\,/\bar{c}\,\underline{p}).
\end{aligned}$$

This shows that the number of events occurring before time T is stochastically dominated by a geometric random variable with parameter $\underline{s}\,/\bar{c}\,\underline{p} > 0$, so that it is finite almost surely and that its expectation is bounded by a constant $\bar{a}$. Since this is true for all particles, we deduce that total number of events has an expectation bounded by $N_{max}\bar{a}$. Finally, observing that this bound does not depend on the initial position, we deduce using the Markov property at times $T, 2T, \ldots$ that the number of event occurring before a time horizon nT is bounded by $kN_{max}\bar{a}$, and hence that it is finite almost surely.

This concludes the proof of Proposition 21. □

3.3 Numerical properties of the constrained branching process

The aim of this section is to discuss some of the numerical properties of the N_{min}–N_{max} process defined in Section 2.1, see Remark 4. In Section 3.3.1, we first make comparisons between this algorithm and the fixed size Moran type algorithm studied in [18, 19]. We will often use the abbreviation FV IPS to refer to the latter, in reference to the name *Fleming Viot interacting particle system*, which has been used in several instances in the literature. In Section 3.3.2, we introduce a filtering method to deal with numerical discrepancies occurring in situations where $\Pi^A_T \Pi^B_T$ has high variance.

3.3.1 Comparison with the FV IPS

We recall that the FV IPS simulates a set of N sub-Markov processes, until the first time one of the particles dies, at which point a resampling event occurs by uniformly selecting one of the remaining $N-1$ particles and duplicating it. Note that the FV IPS is a particular instance of the N_{min}–N_{max} IPS, with a null branching rate and $N_{min} = N_{max} = N$. In particular, Theorem 6 applies to the FV IPS (see [57] for an analogous result in the FV case).

In some situations, it is possible to use the FV IPS to approximate quantities associated with branching processes. Indeed, it is well known that for a branching process $X = (X_t)_{t \geq 0}$, under fairly weak assumptions, the following many-to-one lemma holds

$$\psi_t[f](x) \coloneqq \mathbb{E}_{\delta_x}[\langle f, X_t \rangle] = \mathbf{E}_x\left[\mathrm{e}^{\int_0^t \beta(Y_s)\mathrm{d}s} f(Y_t)\right],$$

where β is determined by the branching rate and mean offspring of X, and the law of Y is uniform amongst all paths of X. In the case where β is uniformly bounded above, one may compensate the right-hand side above by its maximum, $\bar{\beta}$, to introduce a penalised process: $\psi^{\mathtt{k}}_t[f](x) \coloneqq \mathrm{e}^{-\bar{\beta}t}\psi_t[f](x)$. The penalisation has the effect of killing Y at an additional rate $\bar{\beta} - \beta$, thus yielding a sub-Markov process with semigroup $\psi^{\mathtt{k}}_t$. This then allows one to use the

FV IPS to obtain estimates of this semigroup and, in some situations, of its associated quasi-stationary distribution, for example. In general, the use of the FV IPS for the approximation of quantities related to some non-conservative semigroups is widespread, see for instance [6, 18, 51, 32, 56, 34, 8].

In this section, we show on a simple example that the N_{min}–N_{max} algorithm allows one to overcome some of the limitations of the FV IPS. For simplicity, we only consider the situation where $N_{min} = N_{max}$.

A simple branching birth and death process. We consider a birth and death process on $E = \{1, \dots, M\}$ for some parameter $M \in \mathbb{N}$. For a particle occupying state x, one of the following things may occur:

- the particle jumps with rate x^2, to state $1 \vee (x-1)$ with probability $\dfrac{x}{x+1}$ and to state $(x+1) \wedge M$ with probability $\dfrac{1}{x+1}$,
- at rate $b(x) = x$, a new particle is produced at the same site, which will continue independently the same behaviour as the original particle.

In this situation where the state space is finite and irreducible, it is well known that the semigroup $(\psi_t)_{t\geq 0}$ admits a limiting distribution up to an exponential rescaling (see for instance [16]): there exist a constant $\lambda_M \in \mathbb{R}$, a positive function $\eta_M : \{1, \dots, M\} \to (0, +\infty)$ and a probability measure ν_M on $\{1, \dots, M\}$ such that

$$\sup_{f:\{1,\dots,M\},\|f\|_\infty\leq 1} \left|e^{-\lambda_M t}\psi_t[f](x) - \eta_M(x)\nu_M(f)\right| \leq C_M e^{-\gamma_M t}, \qquad t \geq 0,$$

for some positive constants C_M, γ_M.

The associated birth and death process with killing evolves on $\{1, \dots, M\}$ with the same jump rates, without branching, and with the additional killing rate $\kappa(x) = M - x$. Denoting as above by ψ^{k} the associated sub-Markov semigroup, we have $\psi^{\mathrm{k}}_t = e^{-Mt}\psi_t$ and hence

$$\sup_{f:\{1,\dots,M\},\|f\|_\infty\leq 1} \left|e^{-(\lambda_M - M)t}\psi^{\mathrm{k}}_t[f](x) - \eta_M(x)\nu_M(f)\right| \leq C_M e^{-\gamma_M t}.$$

Theoretical convergence of the N_{min}–N_{max} and of the FV IPSs. In what follows, m^M_T refers to the empirical measure of the N_{min}–N_{max} particle system at time T, associated to the above branching birth and death process, and μ^M_T refers to the empirical measure of the FV IPS with $N := N_{min} = N_{max}$ particles at time T associated to the above killed birth and death process.

In particular, Proposition 18 can be adapted to our setting and hence, in the situation where $N_{min} = N_{max} = N$,

$$\left\|\frac{m^M_T(f)}{N} - \nu_M(f)\right\|_2 \leq \frac{(2C_X + C)\|f\|_\infty}{N^{\alpha/2}} + C_X\, e^{-\gamma_X T}\, \|f\|_\infty.$$

for some $\alpha > 0$. Using the fact that the N_{min}–N_{max} process is ergodic, we deduce that its normalised empirical stationary distribution $\mathcal{X}^M_N$ (the normalised empirical distribution of a random vector of particles distributed according to the stationary distribution of the N_{min}–N_{max} process) satisfies

$$\mathcal{X}^M_N(f) \xrightarrow[N\to+\infty]{L_2} \nu_M(f),$$

$N = 10$		$\lvert\mathbb{E}(\hat{\theta}_N) - \nu_M(f_M)\rvert$	$\mathrm{Std}(\hat{\theta}_N)$	$(A_T + B_T)/T$
$M = 10$	N_{min}–N_{max}	0.08	0.30	14.0
	FV	0.10	0.41	87.2
$M = 100$	N_{min}–N_{max}	0.08	0.30	14.0
	FV	0.20	0.51	988
$M = 1000$	N_{min}–N_{max}	0.08	0.30	14.0
	FV	0.22	0.53	9989
$M = +\infty$	N_{min}–N_{max}	0.08	0.30	14.0
	FV	*	*	*

Table 1: For each value of $M \in \{10, 100, 1000, +\infty\}$, we display the bias of the estimator $\hat{\theta}_N$, its standard deviation, and the number of events, for the N_{min}–N_{max} algorithm (for which $\hat{\theta}_N = \mathcal{X}_N^M(f_M)$) and for the FV IPS (for which $\hat{\theta}_N = \mathcal{Y}_N^M(f_M)$) with $N = N_{min} = N_{max} = 10$ particles. Note that the FV IPS is not defined when $M = +\infty$.

for all function $f : \{1, \ldots, M\} \to \nu_M$. Similarly, the normalised empirical stationary distribution $\mathcal{Y}_N^M$ of the FV particle system satisfies

$$\mathcal{Y}_N^M(f) \xrightarrow[N\to+\infty]{L_2} \nu_M(f).$$

These theoretical convergence results ensure that both algorithms can be used to approximate the limiting distribution ν_M. It remains to compare their performances via numerical simulations.

Numerical comparison of the N_{min}–N_{max} and the FV IPSs. We now consider the problem of numerically approximating the mean of ν_M, that is $\nu_M(f_M)$ with $f_M(x) = x$ for all $x \in \{1, \ldots, M\}$. The estimator is $\hat{\theta}_N = \mathcal{X}_N^M(f_M)$ (resp. $\hat{\theta}_N = \mathcal{Y}_N^M(f_M)$) for the N_{min}–N_{max} (resp. FV) IPS, and we compare three important metrics determining the performance of the algorithms:

- the bias of the estimator, $\lvert\mathbb{E}(\hat{\theta}_N) - \nu_M(f_M)\rvert$,
- the standard deviation of the estimator, $\mathrm{Std}(\hat{\theta}_N)$,
- the number of interaction events, $(A_T + B_T)/T$, which corresponds to the number B_T of selection events in the N_{min}–N_{max} algorithm and the number A_T of resampling events in FV IPS, per unit of time.

The last metric $(A_T + B_T)/T$ is particularly important since it entails that the total number of interaction events grows (at least) linearly in time.

For $N = 10$ and then $N = 100$, we let the upper bound of the state space M vary in $\{10, 100, 1000, +\infty\}$, and present the results in Table 1 for $N = 10$, and in Table 2 for $N = 100$. For each metric, a lower number indicates that the algorithm performs better.

$N = 100$		$\lvert\mathbb{E}(\hat{\theta}_N) - \nu_M(f_M)\rvert$	$\mathrm{Std}(\hat{\theta}_N)$	$(A_T + B_T)/T$
$M = 10$	N_{min}–N_{max}	0.01	0.12	144
	FV	0.02	0.18	857
$M = 100$	N_{min}–N_{max}	0.01	0.12	144
	FV	0.10	0.39	9866
$M = 1000$	N_{min}–N_{max}	0.01	0.12	144
	FV	0.20	0.50	99873
$M = +\infty$	N_{min}–N_{max}	0.01	0.12	144
	FV	$*$	$*$	$*$

Table 2: For each value of $M \in \{10, 100, 1000, +\infty\}$, we display the bias of the estimator $\hat{\theta}_N$, its standard deviation, and the number of events, for the N_{min}–N_{max} algorithm (for which $\hat{\theta}_N = \mathcal{X}_N^M(f_M)$) and for the FV IPS (for which $\hat{\theta}_N = \mathcal{Y}_N^M(f_M)$) with $N = N_{min} = N_{max} = 100$ particles. Note that the FV IPS is not defined when $M = +\infty$.

We can see that, for this particular example, the N_{min}–N_{max} algorithm performs better in all situations, with respect to all metrics. Moreover, the bias of the FV approximation increases when M increases, while the variance of the estimator also increases and the computation cost increases. Without surprise, increasing M does not change the dynamic of the N_{min}–N_{max} algorithm, since very few particles, if any, reach the boundary M even when $M = 10$. Further, this suggests that it should be possible to extend the N_{min}–N_{max} algorithm to the case of unbounded branching rates, and this is supported by the numerical simulations when $M = +\infty$ (for which the FV IPS is of course not defined), although it is not covered by our main result. We intend to address these questions in future work.

Remark 22. We note that the birth-death process studied above is a typical example of the type of setting where the N_{min}–N_{max} process may outperform the FV IPS. Here, we briefly outline another example where this is also the case.

We consider a branching random walk on $\{0, \dots, n\}$ for some $n \geq 1$ with a time inhomogeneous branching rate. More precisely, let $(X_t)_{t\geq 0}$ denote a continuous time random walk on $\mathbb{Z}$ that moves one step to the right with probability $p \in (0, 1)$ and one step to the left otherwise. Each step is taken after an (independent) exponentially distributed time with parameter 1.

Now let $b, k : \{0, \dots, n\} \to [0, \infty)$ be bounded measurable functions and $\mathcal{E}_b$, $\mathcal{E}_k$ be exponential random variables, that are independent of each other and X, and define

$$\tau_1^b = \inf\{t > 0 : \int_0^t b(X_s)\mathrm{d}s > \mathcal{E}_b\},$$

and

$$\tau_1^k = \inf\{t > 0 : \int_0^t k(X_s)\mathrm{d}s > \mathcal{E}_k\}.$$

Further, set $\tau_1 = \tau_1^b \wedge \tau_1^k$. If $\tau_1 = \tau_1^b$, a new particle is added to the system at position $X_{\tau_1^b}$, which continues the same behaviour as the original particle. If $\tau_1 = \tau_1^k$, the particle is removed from the system. We can then define τ_2 to be the first time after τ_1 that one of the particles alive at time τ_1 branches or is killed, and this process continues iteratively.

Letting N_t denote the number of particles alive at time $t \geq 0$, and $\{X_t^i : i = 1, \ldots, N_t\}$ their positions, the branching random walk is defined as

$$Z_t = \sum_{i=1}^{N_t} \delta_{X_t^i}, \qquad t \geq 0.$$

We now consider a time-inhomogeneous version of the above process. We consider a sequence of switching times $0 = T_0^{\texttt{off}} < T_1^{\texttt{on}} < T_1^{\texttt{off}} < T_2^{\texttt{on}} < T_2^{\texttt{off}} < \ldots$, defined via

$$T_i^{\texttt{on}} = T_{i-1}^{\texttt{off}} + \mathbf{e}_{i-1}(s_{\texttt{on}}), \qquad i \geq 1,$$

and

$$T_i^{\texttt{off}} = T_i^{\texttt{on}} + \mathbf{e}_i(s_{\texttt{off}}), \qquad i \geq 1,$$

where $\mathbf{e}_i(s_{\texttt{off}})$ and $\mathbf{e}_i(s_{\texttt{on}})$ are independent (of each other and everything else) exponential random variables with rate $s_{\texttt{off}}$ and $s_{\texttt{on}}$, respectively. For $T_{i-1}^{\texttt{off}} \leq t < T_i^{\texttt{on}}$, Z_t evolves as above but with $b = 0$, i.e. each particle moves as a random walk until it is (possibly) killed, until time $T_i^{\texttt{on}}$. At time $T_i^{\texttt{on}}$, we sample from an exponential random variable with rate $B \in (0, \infty)$ and set b to be this value. For $T_i^{\texttt{on}} \leq t < T_i^{\texttt{off}}$, Z_t then evolves as the branching random walk with branching rate b for each particle. At time $T_i^{\texttt{off}}$, we again set the branching rate to be 0, and continue this process until some terminal time T.

We can see that in this case, while the branching rate is bounded almost surely on any compact time interval, we may simulate the process according to the Algorithm 1. However, we cannot use the fixed size Moran type algorithm since we do not know the upper bound for the branching rate.

However, there are other circumstances where the FV IPS may perform at least as well as the N_{min}–N_{max} process due to large deviations effects. We refer the reader to the next section for a discussion and possible solution to this.

3.3.2 Particle filtering methods for the penalised particle system

In many applications, it can be of interest to estimate numerically the growth rate of the semigroup Q for large times. A natural conjecture is that for large times T, $m_0 Q_T f \approx \langle m_0, \psi \rangle \langle \varphi, f \rangle \mathrm{e}^{\lambda T}$, where φ and ψ are the left eigenmeasure and right-eigenfunction respectively of the operator Q, and λ is the *Lyapunov exponent* for Q. One can try to estimate λ by computing the quantity $m_0 Q_T f$ for large times T.

By Theorem 6, one method for estimating λ would be through equation (6), so that we can simulate the (m_T, Π_T^A, Π_T^B)-system and estimate λ by:

$$\hat{\lambda} = \frac{1}{T} \log \left(\Pi_T^A \Pi_T^B m_T(\mathbf{1}) \right) = \frac{1}{T} \log \left(\Pi_T^A \Pi_T^B \right), \tag{33}$$

for T large. In practice, however, this is numerically a poor choice of estimator, due to large-deviation effects. The presence of the 'exponential' functionals, Π_T^A, Π_T^B mean that significant contributions to the expectation will be made by paths of the N_{min}–N_{max} process which have probability $e^{-\delta T}$, and so infeasibly many simulations will be needed to compute an unbiased estimate for $\hat{\lambda}$. A simple solution is to replace the estimator by

$$\bar{\lambda} = \frac{1}{\delta t} \log \left[\frac{1}{n} \sum_{i<n} \left(\frac{\Pi^A_{t_{i+1}} \Pi^B_{t_{i+1}}}{\Pi^A_{t_i} \Pi^B_{t_i}} \right) \right], \tag{34}$$

where $t_i = T\frac{i}{N}$ and $\delta t = \frac{T}{N}$. In many examples, under conditions which ensure that the process is ergodic and thus converges to a stationary distribution, as $T \to \infty$, for fixed δt, the estimator $\bar{\lambda}$ can be shown to converge to λ (e.g. [24]).

In some cases the expression in (34) may not be valid and we can only use estimates of the form in (33). Examples may be when the process is not ergodic, or the system is not time-homogenous, such as the example given in Remark 22. In such circumstances, we propose a novel Monte Carlo method to compute estimates of λ. Specifically, we note that the N_{min}–N_{max} process $\mathcal{X}_t := \{X_t^1, \ldots, X_t^{N_t}\}$ defines a continuous time Markov process, while the processes A_t and B_t corresponding to the number of selection and resampling events, are themselves additive functionals of the Markov process. Estimation of λ corresponds to computing a Feynman-Kac type exponential functional for a Markov process, and it is well known that estimating such functionals can be performed efficiently using interacting particle simulations, [21, 20, 19, 18]. This suggests an algorithm as follows, assuming that b^i, κ^i, p^i and q^i have been chosen:

1. Fix $T, N^*, \delta t, N_0, x_0$. Initiate with $t = 0$, and $\mathcal{X}^1, \ldots, \mathcal{X}^{N^*}$ each containing N_0 particles of initial value x. Set $W = 1$.

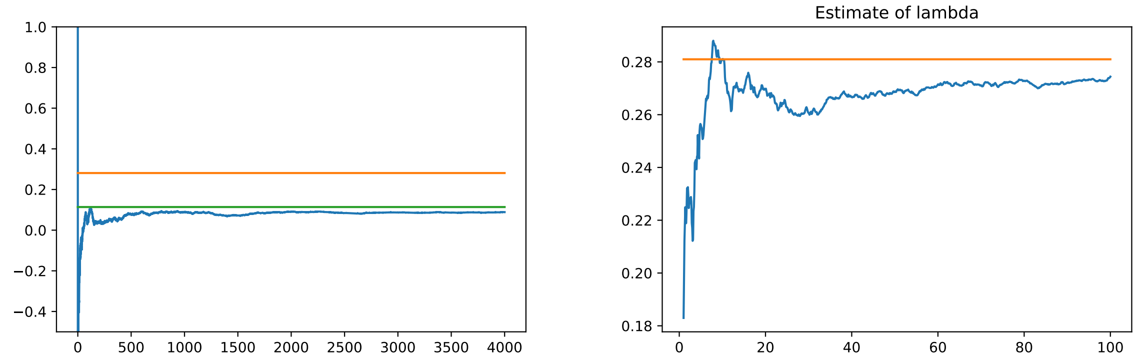
2. Run each $\mathcal{X}^i$ from time t to time $t' = \min\{t + \delta t, T\}$ according to the $N_{min} - N_{max}$ algorithm above, setting $w_i := \frac{\Pi^{A^i}_{t'} \Pi^{B^i}_{t'}}{\Pi^{A^i}_{t} \Pi^{B^i}_{t}}$.

3. Set $W := W \times \left(\frac{1}{N*} \sum_{i=1}^{N^*} w_i \right)$.

4. If $t' < T$, set $t = t'$, resample each $\mathcal{X}^i$ from the existing particles $\{\mathcal{X}^1, \ldots, \mathcal{X}^{N^*}\}$ with probabilities proportional to $w_1, \ldots, w_{N^*}$, independently (multinomial resampling). Return to step 2.

5. If $t' = T$, stop, and return
$$\hat{\lambda} := \frac{1}{T} \log(W).$$

Here, the choice of multinomial resampling could be replaced by other resampling schemes. Similarly, the resampling could be conditional on properties of the particles (e.g. the effective sample size). We compare the effectiveness of the two methods in Figure 2, for a simple Markov process where the true value of the Lyapunov exponent can be calculated accurately numerically. We note that both methods have similar computational effort, and neither appears to converge to the 'true' eigenvalue, however the particle filter method achieves a much better numerical estimate of λ.



(a) Single Trajectory Simulation (b) Particle Filter Simulation

Figure 2: We use two methods described above to estimate the eigenvalue λ. The first simulation (left) estimates λ using (33), based on a single trajectory run up to time $T = 4000$. The true value of λ is highlighted with the orange line, the green line represents an analytical estimate of the convergence value for the simulation (in this case, an estimate of the average growth of the exponential weights under the stationary measure for the system). In the second figure (right), λ is estimated using an interacting particle system, as described above. In this case, the particle filter has 100 particles, and is run for a total time of $T = 40$, giving comparable computational effort. It should be noted however that the second case also has a computational cost associated with the resampling scheme applied, and so the overall computational cost will be higher depending on the resampling scheme applied, and the frequency of resampling events.

4 Proof of Theorem 6

This section is dedicated to the proof of Theorem 6, which is rather long and so we break it up into five steps. In the first three steps, the main idea is to define two martingales $\mathbb{M}$ and $\mathcal{M}$ such that

$$\nu_t^f - \nu_0^f = \Pi_t^A \Pi_t^B (\mathbb{M}_t - \mathbb{M}_{\tau_{C_t}}) + \sum_{n=1}^{A_t} \Pi_{\rho_n}^A \Pi_{\rho_n}^B (\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n -}) + \sum_{n=1}^{B_t} \Pi_{\sigma_n}^A \Pi_{\sigma_n}^B (\mathcal{M}_{\sigma_n} - \mathcal{M}_{\sigma_n -}) + \sum_{n=1}^{C_t} \Pi_{\tau_{n-1}}^A \Pi_{\tau_{n-1}}^B (\mathbb{M}_{\tau_n} - \mathbb{M}_{\tau_{n-1}}), \tag{35}$$

where σ_n and ρ_n denote the n^{th} selection and resampling events, respectively, and

$$\nu_t^f = \Pi_t^A \Pi_t^B \sum_{i \in \bar{S}_t} Q_{T-t} f(X_t^i).$$

The martingale $\mathbb{M}$ compensates the evolution of the branching process between resampling/selection times. The martingale $\mathcal{M}$ compensates the evolution of the particle system during resampling/selection events.

The many-to-one formula (6) then follows. The final two steps make use of the following technical lemma to control the L^2 norms of the martingale increments in order to obtain (7).

Lemma 23. *For all $\ell \geq 1$, for all $T \geq 0$,*

$$\mathbb{E}\left(\left(\Pi_T^B\right)^\ell\right) \leq \exp(c_\ell \|b\|_\infty T),$$

where

$$c_\ell := \sup_{x \geq 2} x\left((1+1/x)^\ell - 1\right).$$

In addition, denoting by β_T the total number of branching events (with or without selection) up to time T, we have

$$\mathbb{E}\left(\beta_T \left(\Pi_T^B\right)^2\right) \leq N_0 \exp((3+c_4)\|b\|_\infty T/2).$$

The above estimates are the main reason why we require b to be bounded. These types of estimates are not difficult to derive in some particular situations where the branching rate is unbounded (typically for time inhomogeneous branching rates, or in situations where the branching rates is high in rarely visited regions), and in these situations, our proofs adapt easily. We defer the proof of the above lemma to the end of the section and now proceed to the proof of Theorem 6.

Proof of Theorem 6.
Step 1. The martingale $\mathbb{M}$. For all $i \in \mathbb{N}$, we define respectively the *birth time* $\mathtt{B}_i$ and the *death time* $\mathtt{D}_i$ of the particle with index i as

$$\mathtt{B}_i = \inf\{t \geq 0,\ i \in \bar{S}_t\} \text{ and } \mathtt{D}_i = \inf\{t \geq \mathtt{B}_i,\ i \notin \bar{S}_t\}.$$

Fix $T > 0$ and $t \in [0, T]$. For all particles $i \in \mathbb{N}$, we set

$$\mathbb{M}_t^{i,n} = \begin{cases} 0, \quad \text{if } t \leq \mathtt{B}_i \vee \tau_n, \\ Q_{T-t} f(X_t^i) - Q_{T-\mathtt{B}_i \vee \tau_n} f(X^i_{\mathtt{B}_i \vee \tau_n}), \quad \text{if } t \in [\mathtt{B}_i \vee \tau_n, \mathtt{D}_i \wedge \tau_{n+1}), \\ (1 - \mathbf{1}_{D^n(i)} + \mathbf{1}_{B^n(i)}) Q_{T-\mathtt{D}_i \wedge \tau_{n+1}} f(X^i_{\mathtt{D}_i \wedge \tau_{n+1}}) - Q_{T-\mathtt{B}_i \vee \tau_n} f(X^i_{\mathtt{B}_i \vee \tau_n}), \quad \text{if } t \geq \mathtt{D}_i \wedge \tau_{n+1}, \end{cases} \tag{36}$$

where $B^n(i)$ denotes the event "τ_{n+1} is a branching time for the i-th particle" and $D^n(i)$ denotes the event "$\tau_{n+1} = \mathtt{D}_i$ is a soft killing time for the i-th particle".

By definition of the process between the time intervals $[\tau_n, \tau_{n+1})$, to show that $\mathbb{M}$ is indeed a martingale, it is sufficient to prove that

$$\mathbb{M}_t^{1,0} := \begin{cases} Q_{T-t} f(X_t^1) - Q_T f(X_0^1), \quad \text{if } t \in [0, \tau_1), \\ (1 - \mathbf{1}_{D(1)} + \mathbf{1}_{B(1)}) Q_{T-\tau_1} f(X_{\tau_1}) - Q_T f(X_0), \quad \text{if } t \geq \tau_1, \end{cases}$$

is a martingale. We observe that, by construction, $\tau_1 = \tau^\partial \wedge \tau^\kappa \wedge \tau^b$ where

$$\begin{aligned} \tau^b &= \inf_{j \in \bar{S}_0} \inf\{t \geq 0,\ \int_0^t b^j((X_u^i)_{i \in \bar{S}_u})\, \mathrm{d}u \geq e^b\}, \\ \tau^\kappa &= \inf_{j \in \bar{S}_0} \inf\{t \geq 0,\ \int_0^t \kappa^j((X_u^i)_{i \in \bar{S}_u})\, \mathrm{d}s \geq e^\kappa\}, \\ \tau^\partial &= \inf_{j \in \bar{S}_0} \inf\{t \geq 0\ X_t^j \in \partial\}. \end{aligned}$$

Then, for all $s,t\geq 0$ such that $s+t\leq T$, setting $\mathbb{E}^t=\mathbb{E}[\cdot\mid\sigma(\bar{S}_u,\,\bar{X}_u,0\leq u\leq t)]$ and using the fact that $Q_u f\equiv 0$ on ∂ for all $u\geq 0$, we have

$$\begin{aligned}\mathbb{E}^t(\mathbb{M}^{1,0}_{t+s})=\mathbb{E}^t&\left[\left(Q_{T-(t+s)}f(X^1_{t+s})-Q_Tf(X_0)\right)\mathbf{1}_{t+s<\tau_1}\right]\\&+\mathbb{E}^t_x\left[\left((1-\mathbf{1}_{D(1)}+\mathbf{1}_{B(1)})Q_{T-\tau_1}f(X_{\tau_1})-Q_Tf(X_0)\right)\mathbf{1}_{t<\tau_1\leq t+s}\right]\\&+\left((1-\mathbf{1}_{D(1)}+\mathbf{1}_{B(1)})Q_{T-\tau_1}f(X_{\tau_1})-Q_Tf(X_0)\right)\mathbf{1}_{\tau_1\leq t}\end{aligned}\tag{37}$$

Using the fact that the particles are independent copies of X before time τ_1, and the Markov property at time t, we obtain

$$\mathbb{E}^t_x\left[Q_{T-(t+s)}f(X^1_{t+s})\mathbf{1}_{t+s<\tau_1}\right]=\mathtt{E}^t\left[Q_{T-(t+s)}f(Y^1_{t+s})e^{-\int_t^{t+s}h_u\,\mathrm{d}u}\mathbf{1}_{t+s<\tau_\partial}\right]\mathbf{1}_{t<\tau_1},$$

where $\mathtt{E}^t$ is the expectation on a probability space Ω' such that $(Y^i_u)_{u\geq t}$, $i\in\bar{S}_0$, are independent copies of X, starting from X^i_t at time t, and where

$$h_u:=\sum_{j\in\bar{S}_0}b^j((Y^i_u)_{i\in\bar{S}_0})+\kappa^j((Y^i_u)_{i\in\bar{S}_0}).$$

Then, using the Markov property at time $t+s$, we obtain

$$\begin{aligned}\mathbb{E}^t_x\left[Q_{T-(t+s)}f(X^1_{t+s})\mathbf{1}_{t+s<\tau_1}\right]&=\mathtt{E}^t\left[e^{\int_{t+s}^T b(Y^1_u)-\kappa(Y^1_u)\,\mathrm{d}u}f(Y^1_T)e^{-\int_t^{t+s}h_u\,\mathrm{d}u}\mathbf{1}_{t+s<\tau_\partial}\right]\mathbf{1}_{t<\tau_1}\\&=\mathtt{E}^t\left[e^{\int_t^T b(Y^1_u)-\kappa(Y^1_u)\,\mathrm{d}u}f(Y^1_T)e^{-\int_t^{t+s}h^1_u\,\mathrm{d}u}\mathbf{1}_{t+s<\tau_\partial}\right]\mathbf{1}_{t<\tau_1},\end{aligned}$$

where $h^1_u:=2b^1((Y^i_u)_{i\in\bar{S}_0})+\sum_{j\in\bar{S}_0,j\neq 1}b^j((Y^i_u)_{i\in\bar{S}_0})+\kappa^j((Y^i_u)_{i\in\bar{S}_0})$. Similarly, since κ^1 is the intensity of the soft killing time for the first particle, and b^1 is the intensity of its branching time, we obtain

$$\begin{aligned}&\mathbb{E}^t_x\left[(1-\mathbf{1}_{D(1)}+\mathbf{1}_{B(1)})Q_{T-\tau_1}f(X^1_{\tau_1})\mathbf{1}_{t<\tau_1\leq t+s}\right]\\&=\mathtt{E}^t\left[\int_t^{(t+s)\wedge\tau_\partial}(h_u-\kappa^1((Y^i_u)_{i\in\bar{S}_0})+b^1((Y^i_u)_{i\in\bar{S}_0}))Q_{T-u}f(Y^1_u)e^{-\int_t^u h_v\,\mathrm{d}v}\mathrm{d}u\right]\mathbf{1}_{t<\tau_1}\\&\quad+\mathtt{E}^t\left[(1-\mathbf{1}_{Y^1_{\tau_\partial}\in\partial})e^{-\int_t^{\tau_\partial}h_u\,\mathrm{d}u}\mathbf{1}_{\tau_\partial\leq t+s}Q_{T-\tau_\partial}f(Y^1_{\tau_\partial})\right]\mathbf{1}_{t<\tau_1}.\end{aligned}$$

Hence, using the fact that Y satisfies the strong Markov property at time τ_∂, and the fact that $Q_{T-\tau_\partial}f(Y^1_{\tau_\partial})=0$ on the event $\{Y^1_{\tau_\partial}\in\partial\}$, we deduce that

$$\begin{aligned}&\mathbb{E}^t_x\left[(1-\mathbf{1}_{D(1)}+\mathbf{1}_{B(1)})Q_{T-\tau_1}f(X^1_{\tau_1})\mathbf{1}_{t<\tau_1\leq t+s}\right]\\&=\mathtt{E}^t\left[\int_t^{(t+s)\wedge\tau_\partial}h^1_u\,e^{\int_u^T(b(Y^1_v)-\kappa(Y^1_v))\,\mathrm{d}v}\,f(Y^1_T)e^{-\int_t^u h_v\,\mathrm{d}v}\,\mathrm{d}u\right]\mathbf{1}_{t<\tau_1}\\&\quad+\mathtt{E}^t\left[e^{-\int_t^{\tau_\partial}h_u\,\mathrm{d}u}\mathbf{1}_{\tau_\partial\leq t+s}e^{\int_{\tau_\partial}^T(b(Y^1_v)-\kappa(Y^1_v))\,\mathrm{d}v}f(Y^1_T)\right]\mathbf{1}_{t<\tau_1}\\&=\mathtt{E}^t\left[e^{\int_t^T(b(Y^1_v)-\kappa(Y^1_v))\,\mathrm{d}v}\,f(Y^1_T)\left(\int_t^{(t+s)\wedge\tau_\partial}h^1_u\,e^{-\int_t^u h^1_v\,\mathrm{d}v}\mathrm{d}u+e^{-\int_t^{\tau_\partial}h^1_v\,\mathrm{d}v}\mathbf{1}_{\tau_\partial\leq t+s}\right)\right]\mathbf{1}_{t<\tau_1},\end{aligned}$$

where we used the assumption $b(Y^1_v) - \kappa(Y^1_v) = b^1((Y^i_u)_{i\in\bar S_0}) - \kappa^1((Y^i_u)_{i\in\bar S_0})$. But

$$e^{-\int_t^{t+s} h^1_u\,\mathrm{d}u}\mathbf{1}_{t+s<\tau_\partial} + \int_t^{(t+s)\wedge\tau_\partial} h^1_u\, e^{-\int_t^u h^1_v\,\mathrm{d}v}\mathrm{d}u + e^{-\int_t^{\tau_\partial} h^1_v\,\mathrm{d}v}\mathbf{1}_{\tau_\partial\le t+s} = 1$$

Hence we deduce from (37) that

$$\begin{aligned}
\mathbb{E}^t_x(\mathbb{M}^{1,0}_{t+s}) &= (1-\mathbf{1}_{D(1)}+\mathbf{1}_{B(1)})\,(Q_{T-\tau_1}f(X_{\tau_1}) - Q_T f(X_0))\,\mathbf{1}_{\tau_1\le t}\\
&\quad + \left(\mathtt{E}^t\left[e^{\int_t^T (b(Y^1_v)-\kappa(Y^1_v))\,\mathrm{d}v}\, f(Y^1_T)\right] - Q_T f(x)\right)\mathbf{1}_{t<\tau_1}\\
&= (1-\mathbf{1}_{D(1)}+\mathbf{1}_{B(1)})Q_{T-\tau_1}f(X_{\tau_1})\mathbf{1}_{\tau_1\le t} + Q_{T-t}f(X^1_t)\,\mathbf{1}_{t<\tau_1} - Q_T f(x)\\
&= \mathbb{M}^{1,0}_t.
\end{aligned}$$

This concludes Step 1.

Step 2. A first martingale decomposition.

Fix $t\ge 0$ and recall that $C_t = \inf\{n\ge 0 : t<\tau_{n+1}\}$. In what follows, we denote by $\rho_1,\rho_2,\dots$ the sequence of times at which a resampling event occurs, and by $\sigma_1,\sigma_2,\dots$ the sequence of times at which a selection event occurs. Note that, if $\tau_n\in\{\rho_1,\rho_2,\dots\}$ or $\tau_n\in\{\sigma_1,\sigma_2,\dots\}$, then $N_n = N_{n-1}\in\{N_{min},\dots,N_{max}\}$.

We will show that, for all $t\ge 0$, we have

$$\begin{aligned}
\mathbb{M}_t := \sum_{n=0}^{C_t}\sum_{i\in\bar S_t}\mathbb{M}^{i,n}_t &= \sum_{i\in\bar S_t} Q_{T-t}f(X^i_t) - \sum_{i\in\bar S_0} Q_T f(X^i_0)\\
&\quad - \sum_{n=1}^{A_t} Q_{T-\rho_n} f(X^{i_n}_{\rho_n}) + \sum_{n=1}^{B_t} Q_{T-\sigma_n} f(X^{j_{n-1}}_{\sigma_n}),
\end{aligned}\tag{38}$$

where we recall that A_t is the number of resampling events up to time t, B_t is the number of selection events up to time t, and where we have set i_n to denote the index of the particle added at time ρ_n and j_{n-1} is the particle killed (due to a selection event) at time σ_n (as above, $X^{j_{n-1}}_{\rho_n}$ denotes the position of the particle before the selection event).

We will show that (38) holds for $n=0,1$, since the general formula then follows using similar arguments.

Noting that for all $t\in[\tau_0,\tau_1)$, $C_t = A_t = B_t = 0$, in this case we have

$$\sum_{n=0}^{C_t}\sum_{i\in\bar S_t}\mathbb{M}^{i,n}_t = \sum_{i\in\bar S_0}\mathbb{M}^{i,0}_t = \sum_{i\in\bar S_0} Q_{T-t}f(X^i_t) - \sum_{i\in\bar S_0} Q_T f(X^i_0),$$

as required.

Let us now describe the evolution of the martingale at time τ_1, depending on the type of event that occurs.

Case $\tau_1 = \rho_1$. In this case, for all $t\in[\tau_1,\tau_2)$, we have $A_t = 1$ and $B_t = 0$, and $\bar S_t =$

$\{1+\max \bar{S}_0\} \cup \bar{S}_0 \setminus \{i'_0\}$, where i'_0 is the particle removed at time τ_1. Thus

$$\begin{aligned}
\sum_{n=0}^{C_t} \sum_{i \in \bar{S}_t} \mathbb{M}_t^{i,n} &= \sum_{i \in \bar{S}_0} \mathbb{M}_t^{i,0} + \sum_{i \in \bar{S}_{\tau_1}} \mathbb{M}_t^{i,1} \\
&= \sum_{i \in \bar{S}_0} \left[(1 - \mathbf{1}_{i=i'_0}) Q_{T-\tau_1} f(X^i_{\tau_1}) - Q_T f(X^i_0)\right] + \sum_{i \in \bar{S}_{\tau_1}} \left[Q_{T-t} f(X^i_t) - Q_{T-\tau_1} f(X^i_{\tau_1})\right] \\
&= \sum_{i \in \bar{S}_{\tau_1}} Q_{T-t} f(X^i_t) - \sum_{i \in \bar{S}_0} Q_T f(X^i_0) - Q_{T-\tau_1} f(X^{1+\max \bar{S}_0}_{\tau_1}) \\
&= \sum_{i \in \bar{S}_t} Q_{T-t} f(X^i_t) - \sum_{i \in \bar{S}_0} Q_T f(X^i_0) - Q_{T-\tau_1} f(X^{i_1}_{\tau_1}),
\end{aligned}$$

as claimed.

Case $\tau_1 = \sigma_1$. Similarly, if $\tau_1 = \sigma_1$, then, for all $t \in [\tau_1, \tau_2)$, we have $A_t = 0$ and $B_t = 1$, and $\bar{S}_t = \{1+\max \bar{S}_0\} \cup \bar{S}_0 \setminus \{j_0\}$. Denoting by j'_0 the particle duplicated at time τ_1, we have

$$\begin{aligned}
\sum_{n=0}^{C_t} \sum_{i \in \bar{S}_t} \mathbb{M}_t^{i,n} &= \sum_{i \in \bar{S}_0} \mathbb{M}_t^{i,0} + \sum_{i \in \bar{S}_{\tau_1}} \mathbb{M}_t^{i,1} \\
&= \sum_{i \in \bar{S}_0} \left[(1 + \mathbf{1}_{i=j'_0}) Q_{T-\tau_1} f(X^i_{\tau_1}) - Q_T f(X^i_0)\right] + \sum_{i \in \bar{S}_{\tau_1}} \left[Q_{T-t} f(X^i_t) - Q_{T-\tau_1} f(X^i_{\tau_1})\right] \\
&= \sum_{i \in S_1} Q_{T-t} f(X^i_t) - \sum_{i \in S_0} Q_T f(X^i_0) + Q_{T-\tau_1} f(X^{j_0}_{\tau_1}) \\
&= \sum_{i \in \bar{S}_t} Q_{T-t} f(X^i_t) - \sum_{i \in \bar{S}_0} Q_T f(X^i_0) + Q_{T-\tau_1} f(X^{j_0}_{\tau_1}).
\end{aligned}$$

Case $\tau_1 \notin \{\rho_1, \sigma_1\}$. If $\tau_1 \notin \{\rho_1, \sigma_1\}$, we have, for all $t \in [\tau_1, \tau_2)$, $A_t = B_t = 0$. If τ_1 is a branching event of the particle with index $j'_0 \in S_0$, then $\bar{S}_t = \{1+\max \bar{S}_0\} \cup \bar{S}_0$

$$\begin{aligned}
\sum_{n=0}^{C_t} \sum_{i \in \bar{S}_t} \mathbb{M}_t^{i,n} &= \sum_{i \in \bar{S}_0} \left[(1 + \mathbf{1}_{i=j'_0}) Q_{T-\tau_1} f(X^i_{\tau_1}) - Q_T f(X^i_0)\right] + \sum_{i \in \bar{S}_{\tau_1}} \left[Q_{T-t} f(X^i_t) - Q_{T-\tau_1} f(X^i_{\tau_1})\right] \\
&= \sum_{i \in \bar{S}_{\tau_1}} Q_{T-t} f(X^i_t) - \sum_{i \in \bar{S}_0} Q_T f(X^i_0),
\end{aligned}$$

where we used the fact that $X^{j'_0}_{\tau_1} = X^{1+\max \bar{S}_0}_{\tau_1}$.

Continuing in this manner, we obtain (38).

Step 3. A second martingale decomposition.

For all time $t \geq 0$ and all bounded measurable function $f : E \to \mathbb{R}$, we denote the integral of f with respect to the occupation measure of the particle system at time t by

$$\mu^f_t := \sum_{i \in \bar{S}_t} Q_{T-t} f(X^i_t). \tag{39}$$

Further, recall the definitions of Π^A_t, Π^B_t and $\nu^f_t = \Pi^A_t \Pi^B_t \mu^f_t$.

In this step, we prove that equation (35) holds with $\mathbb{M}$ defined as in the previous steps, and $\mathcal{M}$ to be defined shortly. To this end, we first note that from Step 2 we obtain

$$\mu_t^f - \mu_0^f = \sum_{n=0}^{C_t} \sum_{i \in \bar{S}_t} \mathbb{M}_t^{i,n} + \sum_{n=1}^{A_t} Q_{T-\rho_n} f(X_{\rho_n}^{i_{n-1}}) - \sum_{n=1}^{B_t} Q_{T-\sigma_n} f(X_{\sigma_n}^{j_{n-1}}). \tag{40}$$

For $n \in \mathbb{N}$, consider the following martingale increments. We define

$$\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-} := Q_{T-\rho_n} f(X_{\rho_n}^{i_{n-1}}) - \frac{1}{\bar{N}_{\rho_n} - 1} \sum_{i \in \bar{S}_{\rho_n-} \setminus \{i'_{n-1}\}} Q_{T-\rho_n} f(X_{\rho_n}^i), \tag{41}$$

where i_{n-1} is the index of the particle duplicated at time ρ_n and i'_{n-1} is the index of the particle removed at time ρ_n (note also that $\bar{N}_{\rho_n} = \bar{N}_{\rho_n-} = |\bar{S}_{\rho_n-}|$). We also define

$$\mathcal{M}_{\sigma_n} - \mathcal{M}_{\sigma_n-} := \frac{1}{\bar{N}_{\sigma_n} + 1} \left(\sum_{i \in \bar{S}_{\sigma_n-}} Q_{T-\sigma_n} f(X_{\sigma_n}^i) + Q_{T-\sigma_n} f(X_{\sigma_n}^{j'_{n-1}}) \right) - Q_{T-\sigma_n} f(X_{\sigma_n}^{j_{n-1}}), \tag{42}$$

where we recall that j_{n-1} is the particle removed at time σ_n and j'_{n-1} is the index of the particle duplicated during the branching event at this time (note also that $\bar{N}_{\sigma_n} = \bar{N}_{\sigma_n-} = |\bar{S}_{\sigma_n-}|$). We assume further that $\mathcal{M}$ is constant except at these jump times.

Then,

$$\begin{aligned} \mu_t^f - \mu_0^f = \mathbb{M}_t + \mathcal{M}_t &+ \sum_{n=1}^{A_t} \frac{1}{\bar{N}_{\rho_n} - 1} \sum_{i \in \bar{S}_{\rho_n-} \setminus \{i'_{n-1}\}} Q_{T-\rho_n} f(X_{\rho_n}^i) \\ &- \sum_{n=1}^{B_t} \frac{1}{\bar{N}_{\sigma_n} + 1} \left(\sum_{i \in \bar{S}_{\sigma_n-}} Q_{T-\sigma_n} f(X_{\sigma_n}^i) + Q_{T-\sigma_n} f(X_{\sigma_n}^{j'_{n-1}}) \right). \end{aligned} \tag{43}$$

For each $n \geq 1$, we define $\mu_{\tau_n-}^f$ and $\nu_{\tau_n-}^f$ by

$$\mu_{\tau_n-}^f = \sum_{i \in S_{n-1}} f(X_{\tau_n}^i) - \mathbf{1}_{\tau_n \in \mathcal{K}} f(X_{\tau_n}^{i'_{n-1}}) + \mathbf{1}_{\tau_n \in \mathcal{B}} f(X_{\tau_n}^{j'_{n-1}}) \text{ and } \nu_{\tau_n-} = \Pi_{\tau_{n-1}}^A \Pi_{\tau_{n-1}}^B \mu_{\tau_n-}$$

where $\mathcal{K}$ is the set of killing times and $\mathcal{B}$ is the set of branching times. Informally, μ_{τ_n-} and ν_{τ_n-} represent the state of the particle system at time τ_n before the resampling or selection eventually occurring at time τ_n. Note that, at any time τ_n that is not a resampling or a selection time, we have $\mu_{\tau_n-} = \mu_{\tau_n}$ and $\nu_{\tau_n-} = \nu_{\tau_n}$. Hence, denoting by $(\theta_n)_n$ the sequence of events from the families ρ or σ, we obtain for all $t \geq 0$,

$$\begin{aligned} \nu_t^f - \nu_0^f &= \nu_t - \nu_{\tau_{C_t}} + \sum_{n=1}^{C_t} (\nu_{\tau_n} - \nu_{\tau_n-}) + \sum_{n=1}^{C_t} \left(\nu_{\tau_n-} - \nu_{\tau_{n-1}} \right) \\ &= \Pi_t^A \Pi_t^B (\mu_t - \mu_{\tau_{C_t}}) + \sum_{n=1}^{A_t} \nu_{\rho_n} - \nu_{\rho_n-} + \sum_{n=1}^{B_t} \nu_{\sigma_n} - \nu_{\sigma_n-} + \sum_{n=1}^{C_t} \nu_{\tau_n-} - \nu_{\tau_{n-1}}. \end{aligned}$$

At resampling times ρ_n, we have $\Pi^B_{\rho_n} - \Pi^B_{\rho_n-} = 0$ and hence,

$$\begin{aligned}
\nu^f_{\rho_n} - \nu^f_{\rho_n-} &= \Pi^A_{\rho_n}\Pi^B_{\rho_n}\mu^f_{\rho_n} - \Pi^A_{\rho_{n-1}}\Pi^B_{\rho_{n-1}}\mu^f_{\rho_n-} \\
&= \Pi^A_{\rho_n}\Pi^B_{\rho_n}(\mu^f_{\rho_n} - \mu^f_{\rho_n-}) + \Pi^B_{\rho_n}\mu^f_{\rho_n-}\left(\Pi^A_{\rho_n} - \Pi^A_{\rho_{n-1}}\right) \\
&= \Pi^A_{\rho_n}\Pi^B_{\rho_n}\, Q_{T-\rho_n} f(X^{i_{n-1}}_{\rho_n}) + \Pi^B_{\rho_n}\mu^f_{\rho_n-}\left(\Pi^A_{\rho_n} - \Pi^A_{\rho_{n-1}}\right) \qquad (44)
\end{aligned}$$

For the increment of Π^A, we note that

$$\Pi^A_{\rho_n} - \Pi^A_{\rho_{n-1}} = \Pi^A_{\rho_n} - \left(\frac{\bar{N}_{\rho_n}}{\bar{N}_{\rho_n} - 1}\right)\Pi^A_{\rho_n} = -\frac{1}{\bar{N}_{\rho_n} - 1}\Pi^A_{\rho_n}.$$

Substituting this into (44) and using (41), we obtain

$$\begin{aligned}
\nu^f_{\rho_n} - \nu^f_{\rho_n-} &= \Pi^A_{\rho_n}\Pi^B_{\rho_n}\left(Q_{T-\rho_n} f(X^{i_{n-1}}_{\rho_n}) - \frac{1}{\bar{N}_{\rho_n} - 1}\mu^f_{\rho_n-}\right) \\
&= \Pi^A_{\rho_n}\Pi^B_{\rho_n}\left(\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-}\right). \qquad (45)
\end{aligned}$$

Similarly, for $n \in \{1, \dots, B_t\}$,

$$\nu^f_{\sigma_n} - \nu^f_{\sigma_n-} = \Pi^A_{\sigma_n}\Pi^B_{\sigma_n}(\mu^f_{\sigma_n} - \mu^f_{\sigma_n-}) + \Pi^A_{\sigma_n}\mu^f_{\sigma_n-}\left(\Pi^B_{\sigma_n} - \Pi^B_{\sigma_n-}\right). \qquad (46)$$

In this case, we have

$$\Pi^B_{\sigma_n} - \Pi^B_{\sigma_n-} = \Pi^B_{\sigma_n} - \left(\frac{\bar{N}_{\sigma_n}}{\bar{N}_{\sigma_n} + 1}\right)\Pi^B_{\sigma_n} = \frac{1}{\bar{N}_{\sigma_n} + 1}\Pi^B_{\sigma_n}, \qquad (47)$$

and

$$\mu^f_{\sigma_n} - \mu^f_{\sigma_n-} = -Q_{T-\sigma_n} f(X^{j_{n-1}}_{\sigma_n}), \qquad (48)$$

where j_{n-1} is the index of the particle removed from the system at time σ_n. Again, substituting these equalities back into (46) and using (42) gives,

$$\begin{aligned}
\nu^f_{\sigma_n} - \nu^f_{\sigma_n-} &= \Pi^A_{\sigma_n}\Pi^B_{\sigma_n}\left(-Q_{T-\sigma_n} f(X^{j_{n-1}}_{\sigma_n}) + \frac{1}{\bar{N}_{\sigma_n} + 1}\mu^f_{\sigma_n-}\right) \\
&= \Pi^A_{\sigma_n}\Pi^B_{\sigma_n}(\mathcal{M}_{\sigma_n} - \mathcal{M}_{\sigma_n-}). \qquad (49)
\end{aligned}$$

Finally, we have $\nu_{\tau_n-} - \nu_{\tau_{n-1}} = \Pi^A_{\tau_{n-1}}\Pi^B_{\tau_{n-1}}(\mathbb{M}_{\tau_n} - \mathbb{M}_{\tau_{n-1}})$ for all $n \geq 1$, hence

$$\begin{aligned}
\nu^f_t - \nu^f_0 = \Pi^A_t\Pi^B_t(\mathbb{M}_t - \mathbb{M}_{\tau_{C_t}}) + \sum_{n=1}^{A_t}\Pi^A_{\rho_n}\Pi^B_{\rho_n}\left(\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-}\right) + \sum_{n=1}^{B_t}\Pi^A_{\sigma_n}\Pi^B_{\sigma_n}(\mathcal{M}_{\sigma_n} - \mathcal{M}_{\sigma_n-}) \\
+ \sum_{n=1}^{C_t}\Pi^A_{\tau_{n-1}}\Pi^B_{\tau_{n-1}}(\mathbb{M}_{\tau_n} - \mathbb{M}_{\tau_{n-1}}),
\end{aligned}$$

which entails (35) and thus concludes Step 3.

Step 4. Control of the L^2 norm of the martingale increments. In this section, we assume without loss of generality (taking $e^{-\|b\|_\infty T} f/\|f\|_\infty$ instead of f) that

$$\sup_{t\in[0,T]} \|Q_t f\|_\infty \leq 1.$$

We now show that the L^2 norm of each of these quantities on the righthand side of (35) is finite. To this end, let us first consider $\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-}$. We have, according to (41),

$$|\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-}| = \left| Q_{T-\rho_n} f(X_{\rho_n}^{i_{n-1}}) - \frac{1}{\bar{N}_{\rho_n} - 1} \sum_{i \in \bar{S}_{\rho_n-} \setminus \{i'_{n-1}\}} Q_{T-\rho_n} f(X_{\rho_n}^i) \right| \leq 2. \tag{50}$$

Then, setting $R_T = \max_{t \in [0,T]} N_t$, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{n=1}^{A_T} \Pi_{\rho_n}^A \Pi_{\rho_n}^B \left(\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-}\right)\right)^2\right] \\
&\qquad = \mathbb{E}\left[\left(\frac{\bar{N}_{\rho_n} - 1}{\bar{N}_{\rho_n}} \sum_{n=1}^{A_T} \Pi_{\rho_n-}^A \Pi_{\rho_n-}^B \left(\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-}\right)\right)^2\right] \\
&\qquad \leq 4\, \mathbb{E}\left[\sum_{n=1}^{A_T} \left(\frac{R_T - 1}{R_T}\right)^{2A_{\rho_n}} (\Pi_{\rho_n-}^B)^2\right] \\
&\qquad \leq 4\, \mathbb{E}\left[(\Pi_T^B)^2 \sum_{n=1}^{A_T} \left(\frac{R_T - 1}{R_T}\right)^{2n}\right].
\end{aligned} \tag{51}$$

To control the sum term, since $1/R_T \leq \dfrac{1}{2}$, we have

$$\begin{aligned}
\sum_{n=1}^{A_T} \left(1 - \frac{1}{R_T}\right)^{2n} \leq \sum_{n=1}^{+\infty} \left(1 - \frac{1}{R_T}\right)^{2n} &= \frac{(1 - \frac{1}{R_T})^2}{1 - (1 - \frac{1}{R_T})^2} \\
&= R_T \frac{1 - 2/R_T + R_T^{-2}}{2 - \frac{1}{R_T}} \leq R_T/2 \leq (N_0 + \beta_T)/2,
\end{aligned}$$

where β_T denotes the total number of branching events before time T. We now make use of Lemma 23, stated and proved at the end of this section, which entails that

$$\mathbb{E}\left[(\Pi_T^B)^2 (N_0 + \beta_T)/2\right] \leq N_0 \exp(c\|b\|_\infty T), \tag{52}$$

for some (explicit) constant $c > 0$. Substituting these estimates into (51), we obtain

$$\mathbb{E}\left[\left(\sum_{n=1}^{A_T} \Pi_{\rho_n}^A \Pi_{\rho_n}^B \left(\mathcal{M}_{\rho_n} - \mathcal{M}_{\rho_n-}\right)\right)^2\right] \leq 4\, N_0 \, \exp\left(c\|b\|_\infty T\right). \tag{53}$$

Let us now consider $\Delta_{\sigma_n} \mathcal{M}$. Using (42), we have

$$|\mathcal{M}_{\sigma_n} - \mathcal{M}_{\sigma_n-}| \leq 2.$$

Then

$$\mathbb{E}\left[\left(\sum_{n=1}^{B_T}\Pi^A_{\sigma_n}\Pi^B_{\sigma_n}(\mathcal{M}_{\sigma_n}-\mathcal{M}_{\sigma_{n-}})\right)^2\right] \tag{54}$$

$$\leq \mathbb{E}\left[\left(\sum_{n=1}^{B_T}\Pi^A_{\sigma_{n-}}\Pi^B_{\sigma_{n-}}\,\frac{N_{\sigma_{n-}}+1}{N_{\sigma_{n-}}}(\mathcal{M}_{\sigma_n}-\mathcal{M}_{\sigma_{n-}})\right)^2\right]$$

$$\leq 4\,\mathbb{E}\left[\sum_{n=1}^{B_T}\left(\Pi^B_{\sigma_n}\right)^2\right] \leq 4\,\mathbb{E}\left[\beta_T\,\left(\Pi^B_T\right)^2\right]. \tag{55}$$

Using Lemma 23 stated and proved at the end of this section, we deduce that

$$\mathbb{E}\left[\left(\sum_{n=1}^{B_T}\Pi^A_{\sigma_n}\Pi^B_{\sigma_n}(\mathcal{M}_{\sigma_n}-\mathcal{M}_{\sigma_{n-}})\right)^2\right] \leq 4\,N_0\exp(c\|b\|_\infty T), \tag{56}$$

where c can be chosen to be the same as in (53).

It remains to control the increments of $\mathbb{M}$. For all $n\geq 1$, let $E_n\subset\mathbb{N}$ be the set of indices of particles defined by

$$E_n=\{i\in\mathbb{N}\text{ s.t. the }i^{th}\text{ particle is branching, killed, born, removed or duplicated at time }\tau_n\}.$$

Observe that $|E_n|\leq 4$ for all $n\geq 1$. We also define, for all $i\in\mathbb{N}$, the sequence of times $(\tau^i_n)_{n\geq 0}$ by $\tau^i_0=0$ and

$$\tau^i_{n+1}=\min\{\tau_k>\tau^i_n,\text{ with }k\geq 0\text{ and }i\in E_k\}.$$

We write

$$\mathbb{M}_T=\sum_{i\in\mathbb{N}}\sum_{n\geq 1}\mathbb{M}^i_{\tau^i_n\wedge T}-\mathbb{M}^i_{\tau^i_{n-1}\wedge T}.$$

Since all the martingales increments are orthogonal, we have, for all $i\in\mathbb{N}$,

$$\Delta := \mathbb{E}\left[\left(\Pi^A_T\Pi^B_T(\mathbb{M}_T-\mathbb{M}_{\tau_{C_T}})+\sum_{n=1}^{C_T}\Pi^A_{\tau_{n-1}}\Pi^B_{\tau_{n-1}}(\mathbb{M}_{\tau_n}-\mathbb{M}_{\tau_{n-1}})\right)^2\right]$$

$$=\mathbb{E}\left[\left(\sum_{i\in\mathbb{N}}\sum_{n\geq 1}\Pi^A_{\tau^i_{n-1}}\Pi^B_{\tau^i_{n-1}}(\mathbb{M}^i_{\tau^i_n\wedge T}-\mathbb{M}^i_{\tau^i_{n-1}\wedge T})\right)^2\right]$$

$$=\mathbb{E}\left[\sum_{i\in\mathbb{N}}\sum_{n\geq 1}\left(\Pi^A_{\tau^i_{n-1}}\Pi^B_{\tau^i_{n-1}}\right)^2\left(\mathbb{M}^i_{\tau^i_n\wedge T}-\mathbb{M}^i_{\tau^i_{n-1}\wedge T}\right)^2\right].$$

But the increments $\mathbb{M}^i_{\tau^i_{n+1}\wedge T} - \mathbb{M}^i_{\tau^i_n\wedge T}$ are almost surely bounded by 3, so that

$$\begin{aligned}\Delta &\leq 9\,\mathbb{E}\left[\sum_{i\in\mathbb{N}}\sum_{n\geq 1}\left(\Pi^A_{\tau^i_{n-1}}\Pi^B_{\tau^i_{n-1}}\right)^2\mathbf{1}_{\tau^i_{n-1}<T}\right]\\ &\leq 9\,\mathbb{E}\left[\sum_{n\geq 1}\sum_{i\in E_n}\left(\Pi^A_{\tau_{n-1}}\Pi^B_{\tau_{n-1}}\right)^2\mathbf{1}_{\tau_{n-1}<T}\right]\\ &\leq 36\,\mathbb{E}\left[\sum_{n\geq 1}\left(\Pi^A_{\tau_{n-1}}\Pi^B_{\tau_{n-1}}\right)^2\mathbf{1}_{\tau_{n-1}<T}\right]\end{aligned}$$

Hence, denoting by β_T (resp. κ_T) the number of branching (resp. killing without resampling) events occuring before time $T\geq 0$ (so that $C_T=\beta_T+\kappa_T+A_T$), we have

$$\Delta \leq 36\,\mathbb{E}\left[\left(\Pi^B_T\right)^2\left(1+\beta_T+\kappa_T+\sum_{n=1}^{A_T}\left(\Pi^A_{\rho_{n-1}}\right)^2\right)\right] \tag{57}$$

We already computed that, almost surely,

$$\sum_{n=1}^{A_T}\left(\Pi^A_{\rho_{n-1}}\right)^2 \leq \sum_{k=0}^{+\infty}\left(\frac{R_T-1}{R_T}\right)^{2k} \leq R_T/2 \leq (N_0+\beta_T)/2.$$

Now, observe that the number of particles in the system as time T is upper bounded by $N_0+\beta_T-\kappa_T$ and lower bounded by 0, so that $\kappa_T\leq N_0+\beta_T$, and hence

$$\mathbb{E}\left[\left(\Pi^B_T\right)^2\left(1+\beta_T+\kappa_T+\sum_{n=1}^{A_T}\left(\Pi^A_{\rho_{n-1}}\right)^2\right)\right] \;\leq\; \mathbb{E}\left[\left(\Pi^B_T\right)^2\left(\beta_T+3N_0\right)\right] \;\leq\; 4N_0\exp(c\|b\|_\infty T),$$

where we used, as above, Lemma 23. We thus obtained, using the last inequality and (57), that

$$\Delta \leq 144\,N_0\,\exp\left(c\|b\|_\infty T\right).$$

Using (35), (53), (56) and the last inequality, we deduce that

$$\mathbb{E}\left[\left(\nu^f_T-\nu^f_0\right)^2\right] \leq 152N_0\exp\left(c\|b\|_\infty T\right).$$

Without assuming that $\sup_{t\in[0,T]}\|Q_tf\|_\infty\leq 1$, we thus obtain

$$\mathbb{E}\left[\left(\nu^f_T-\nu^f_0\right)^2\right] \leq 152N_0\exp\left(c\|b\|_\infty T\right)\left(\sup_{t\in[0,T]}\|Q_tf\|_\infty\right)^2. \tag{58}$$

Step 5. Conclusion.

Recall the occupation measure m_t defined in (5) for all $t\geq 0$. With this notation, the conclusion of Step 4 reads

$$\mathbb{E}\left[\left(m_0Q_Tf-\Pi^A_T\Pi^B_Tm_Tf\right)^2\right] \leq 152N_0\exp\left(c\|b\|_\infty T\right)\left(\sup_{t\in[0,T]}\|Q_tf\|_\infty\right)^2.$$

We deduce that

$$\begin{aligned}
&\left\|m_0 Q_T \mathbf{1}_E \,\frac{m_T f}{m_T \mathbf{1}_E} \mathbf{1}_{m_T\neq 0} - m_0 Q_T f\right\|_2 \\
&\qquad \leq \left\|\left(m_0 Q_T \mathbf{1}_E - \Pi^A_T \Pi^B_T m_T \mathbf{1}_E\right) \frac{m_T f}{m_T \mathbf{1}_E} \mathbf{1}_{m_T\neq 0}\right\|_2 + \left\|\Pi^A_T \Pi^B_T m_T f - m_0 Q_T f\right\|_2 \\
&\qquad \leq 2\sqrt{152 N_0}\, \exp\left(c\|b\|_\infty T/2\right) \|f\|_\infty \sup_{t\in[0,T]} \|Q_t \mathbf{1}_E\|_\infty.
\end{aligned}$$

We conclude that

$$\left\|\frac{m_0 Q_T f}{m_0 Q_T \mathbf{1}_E} - \frac{m_T f}{m_T(\mathbf{1}_E)} \mathbf{1}_{m_T\neq 0}\right\|_2 \leq \frac{2\sqrt{152}\exp\left(c\|b\|_\infty T/2\right)\|f\|_\infty \sup_{t\in[0,T]}\|Q_t \mathbf{1}_E\|_\infty}{m_0 Q_T \mathbf{1}_E / N_0} \frac{1}{\sqrt{N_0}}.$$

Observing that $\sup_{t\in[0,T]} \|Q_t \mathbf{1}_E\|_\infty \leq e^{\|b\|_\infty T}$, this concludes the proof of Theorem 6. □

Proof of Lemma 23. We use the same notations as in the proof of Theorem 6.

Fix $C > c_\ell \|b\|_\infty$ and $\delta > 0$ such that $C - \delta C^2 2^\ell \geq c_\ell \|b\|_\infty$. We have

$$\begin{aligned}
\mathbb{E}\left(\left(\Pi^B_{\tau_1\wedge\delta}\right)^\ell e^{-C(\tau_1\wedge\delta)}\right) &= \mathbb{E}\left(\left(\frac{N_0+1}{N_0}\right)^\ell \mathbf{1}_{\sigma_1=\tau_1\wedge\delta} e^{-C(\tau_1\wedge\delta)}\right) + \mathbb{E}\left(\mathbf{1}_{\sigma_1>\tau_1\wedge\delta} e^{-C(\tau_1\wedge\delta)}\right) \\
&= \left(\frac{N_0+1}{N_0}\right)^\ell \mathbb{E}\left(e^{-C(\tau_1\wedge\delta)}\right) + \left(1 - \left(\frac{N_0+1}{N_0}\right)^\ell\right) \mathbb{E}\left(\mathbf{1}_{\sigma_1>\tau_1\wedge\delta} e^{-C(\tau_1\wedge\delta)}\right).
\end{aligned}$$

On the one hand, we have

$$\mathbb{E}\left(e^{-C(\tau_1\wedge\delta)}\right) \leq 1 - e^{-C\delta} C\mathbb{E}(\tau_1\wedge\delta) \leq 1 - (C - \delta C^2)\mathbb{E}(\tau_1\wedge\delta),$$

and, on the other hand,

$$\mathbb{E}\left(\mathbf{1}_{\sigma_1>\tau_1\wedge\delta} e^{-C(\tau_1\wedge\delta)}\right) \geq \mathbb{E}\left(e^{-(\tau_1\wedge\delta)(N_0\|b\|_\infty + C)}\right) \geq 1 - (N_0\|b\|_\infty + C)\mathbb{E}\left(\tau_1\wedge\delta\right).$$

Setting, for the sake of readability, $\alpha = \mathbb{E}(\tau_1\wedge\delta)$ and $\beta = \left(\frac{N_0+1}{N_0}\right)^\ell$, we deduce that

$$\begin{aligned}
\mathbb{E}\left(\left(\Pi^B_{\tau_1\wedge\delta}\right)^\ell e^{-C(\tau_1\wedge\delta)}\right) &\leq 1 - \alpha\left(C - \delta C^2\beta + N_0\|b\|_\infty(1-\beta)\right) \\
&\leq 1 - \alpha\left(C - \delta C^2 2^\ell - \|b\|_\infty c_\ell\right) \leq 1,
\end{aligned} \tag{59}$$

where we used the fact that $\beta \leq 2^\ell$ and $\beta - 1 \leq c_\ell/N_0$.

We define the sequence of stopping times $(\theta_n)_{n\geq 0}$ by $\theta_0 = 0$ and

$$\theta_{n+1} := \inf\{((n+1)\delta)\wedge\tau_k,\ k\geq 1, \tau_k > \theta_n\}.$$

Then

$$\begin{aligned}
\mathbb{E}\left(\left(\Pi^B_{T\wedge\theta_{n+1}}\right)^\ell e^{-C(T\wedge\theta_{n+1})} \mid \mathcal{F}_{T\wedge\theta_n}\right) &= \left(\Pi^B_{T\wedge\theta_n}\right)^\ell e^{-C(T\wedge\theta_n)} \\
&\qquad \times \mathbb{E}_{\mathbb{X}_{T\wedge\theta_n}}\left(\left(\Pi^B_{(T-u)\wedge\theta_1}\right)^\ell e^{-C((T-u)\wedge\theta_1)}\right)_{|u=T\wedge\theta_n}.
\end{aligned}$$

Since $T-u\leq\delta$ almost surely in the right hand term, we deduce from (59) that

$$\mathbb{E}\left(\left(\Pi^B_{T\wedge\theta_{n+1}}\right)^\ell e^{-C(T\wedge\theta_{n+1})}\mid\mathcal{F}_{T\wedge\theta_n}\right)\leq\left(\Pi^B_{T\wedge\theta_n}\right)^\ell e^{-C(T\wedge\theta_n)}.$$

In particular, using Fatou's Lemma, we deduce, letting $n\to+\infty$, that

$$\mathbb{E}\left(\left(\Pi^B_T\right)^\ell e^{-CT}\right)\leq 1.$$

Since this is true for any $C>c_\ell\|b\|_\infty$, this concludes the proof of the first part of the lemma.

For the second part, we observe that β_T is stochastically dominated by a continuous time counting process $(Z_t)_{t\geq 0}$ on $\mathbb{N}$ starting from N_0 and with increment rate $\|b\|_\infty n$ from state n to state $n+1$. In particular, denoting by L the infinitesimal generator of this process and $V(n)=n^2$ for all $n\geq 1$, we obtain

$$LV(n)=\|b\|_\infty n((n+1)^2-n^2)=\|b\|_\infty n(2n+1)\leq 3\|b\|_\infty V(n),$$

and deduce that

$$\mathbb{E}(\beta_T^2)\leq\mathbb{E}(Z_T^2)\leq e^{3\|b\|_\infty T}Z_0^2=e^{3\|b\|_\infty T}N_0^2.$$

Hence, we obtain

$$\mathbb{E}\left(\beta_T\left(\Pi^B_T\right)^2\right)\leq\sqrt{\mathbb{E}(\beta_T^2)}\sqrt{\mathbb{E}\left(\left(\Pi^B_T\right)^4\right)}\leq N_0\exp((3+c_4)\|b\|_\infty T/2).$$

This concludes the proof of the second part of the Lemma. □

Remark 24. Our proof adapts to the case where $\tau_\infty<+\infty$ with positive probability (that is when Assumption 2 does not holds true), using the following strategy. In this situation, one shows using the same approach that, for all $n\geq 0$,

$$m_0Q_Tf=\mathbb{E}\left(\Pi^A_{\tau_n\wedge T}\,\Pi^B_{\tau_n\wedge T}m_{\tau_n\wedge T}Q_{T-\tau_n\wedge T}f\right).$$

Since the branching rate is uniformly bounded and since $\Pi^A_{\tau_n\wedge T}$ is bounded by 1, one checks that $\Pi^A_{\tau_n\wedge T}\,\Pi^B_{\tau_n\wedge T}m_{\tau_n\wedge T}Q_{T-\tau_n\wedge T}f$ can be uniformly bounded (in n) by an integrable random variable. Since in addition $\Pi^A_{\tau_n\wedge T}\,\Pi^B_{\tau_n\wedge T}m_{\tau_n\wedge T}Q_{T-\tau_n\wedge T}f\to 0$ on the event $\tau_\infty\leq T$ (since this events necessarily corresponds to the accumulation of resampling events) and $\Pi^A_{\tau_n\wedge T}\,\Pi^B_{\tau_n\wedge T}m_{\tau_n\wedge T}Q_{T-\tau_n\wedge T}f\to\Pi^A_T\,\Pi^B_Tm_Tf$ on the complementary event, one deduces from the dominated convergence theorem that

$$m_0Q_Tf=\mathbb{E}\left(\Pi^A_T\,\Pi^B_Tm_Tf\mathbf{1}_{T<\tau_\infty}\right).$$

The rest of the proof (see Steps 4 and 5) then proceeds almost identically.

Remark 25. We comment on some direct and natural continuations of the proof, which are easily derived from the fact that the proof is based on the control of martingale increments.

First, if the martingale $\nu_t^f-\nu_0^f$ is càdlàg, then Doob's maximal inequality and (58) entail that

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left(\nu_T^f-\nu_0^f\right)^2\right]\leq c_T\|f\|_\infty^2\sup_{t\in[0,T]}\|Q_t\mathbf{1}_E\|_\infty^2N_0$$

for some (explicit) constant c_T. Then the conclusion reads

$$\left\| \sup_{t\in[0,T]} \left| \frac{m_0 Q_t f}{m_0 Q_t \mathbf{1}_E} - \frac{m_t f}{m_t(\mathbf{1}_E)} \mathbf{1}_{m_t\neq 0} \right| \right\|_2 \leq \frac{c_T \|f\|_\infty}{m_0 Q_T \mathbf{1}_E / N_0} \frac{1}{\sqrt{N_0}},$$

Second, if the killing rate is bounded, then the statement of Lemma 23 holds true with β_T replaced by the total number of killing/branching events C_T, and with different constants. Then, for any $p \geq 2$, using the same approach as in Step 4 of the proof, but using the Burkholder-Davis-Gundy inequality instead, one deduces similarly that

$$\mathbb{E}\left[\left| \nu_T^f - \nu_0^f \right|^p \right] \leq c_T N_0^{p/2},$$

and hence

$$\left\| \frac{m_0 Q_t f}{m_0 Q_t \mathbf{1}_E} - \frac{m_t f}{m_t(\mathbf{1}_E)} \right\|_p \leq \frac{c'_T \|f\|_\infty}{m_0 Q_T \mathbf{1}_E / N_0} \frac{1}{\sqrt{N_0}},$$

for some (explicit) constants c_T and c'_T. Note that the assumption that the killing rate is uniformly bounded is technical: after using Burkholder-Davis-Gundy inequality, one needs to control the term $\mathbb{E}\left(C_T^{p/2} \left(\frac{N+1}{N} \right)^{pC_T} \right)$, however, depending on the particular process at hand, there may be other ways to obtain upper bounds on this quantity which do not require the killing rate to be bounded.

A Formal construction of the particle system

In this section, we provide a formal construction of the particle system described above, and show that it defines a Markov process with respect to its natural filtration.

We consider the measurable space

$$F = \bigcup_{s\in\mathcal{P}_f(\mathbb{N})} \{s\} \otimes \Omega^s \otimes \mathbb{R}_+^s \otimes \mathbb{R}_+^s,$$

where, as usual, denumerable spaces are endowed with the discrete σ-fields, $\mathbb{R}_+$ is endowed with its Borel σ-field, and product spaces are endowed with the product σ-field. We denote generic elements of F by

$$\theta = (s, (\omega_i)_{i\in s}, (e_i^b)_{i\in s}, (e_i^\kappa)_{i\in s}),$$

where $s \in \mathcal{P}_f(\mathbb{N})$ is used to enumerate the particle system, where, for all $i \in s$, $X(\omega_i) : t \in [0, +\infty) \to X_t(\omega_i) \in E \cup \partial$ is the trajectory of the i^{th} particle, and where $e_i^b \in \mathbb{R}_+$ and $e_i^\kappa \in \mathbb{R}_+$ are auxiliary components used to define the branching and soft-killing times.

Fix $s \in \mathcal{P}_f(\mathbb{N})$ and let us consider the probability kernel N^1 from $\{s\} \otimes E^s$ to F defined by

$$N^1(s, (x_i)_{i\in s}) := \delta_s \otimes \left(\bigotimes_{i\in s} \mathbf{P}_{x_i} \right) \otimes \mathcal{E}^{\otimes s} \otimes \mathcal{E}^{\otimes s},$$

where $\mathcal{E}$ denotes the exponential distribution with parameter 1. This kernel is used to model the evolution of the particles before the occurence of a branching/killing event.

Given an element $\theta \in F$, our aim is now to define the position of the system at the first time of branching/killing. In order to do so, we define, for all $\theta \in F$ and all $i \in s$, the functions

$$\begin{aligned}
\tau^{i,b}(\theta) &= \inf\{t \geq 0,\ \int_0^t b^{i_0}((X_u(\omega_j))_{j\in s})\,\mathrm{d}u = e_i^b\},\\
\tau^{i,\kappa}(\theta) &= \inf\{t \geq 0,\ \int_0^t \kappa^{i_0}((X_u(\omega_j))_{j\in s})\,\mathrm{d}u = e_i^\kappa\},\\
\tau^{i,\partial}(\theta) &= \inf\{t \geq 0,\ X_t(\omega_i) \in \partial\}.
\end{aligned}$$

Note that, for all $t \geq 0$, $(u,\omega) \in [0,+\infty)\times\mathcal{F} \mapsto X_u(\omega)$ is measurable with respect $\mathcal{B}([0,+\infty))\otimes\mathcal{F}$, which entails that $\omega \mapsto \int_0^t b^i((X_u(\omega_j))_{j\in s})\,\mathrm{d}u$ is $\mathcal{F}^{\otimes s}$-measurable. In addition, since $t \mapsto \int_0^t b^i((X_u(\omega_j))_{j\in s})\,\mathrm{d}u$, $t \mapsto \int_0^t \kappa^i((X_u(\omega_j))_{j\in s})\,\mathrm{d}u$ and $t \mapsto \mathbf{1}_{X_t(\omega_i)\in\partial}$ are right-continuous, the infima in the above definitions can be taken over $\mathbb{Q}$ and are thus measurable functions of $(\omega_j)_{j\in s}$, $(e_j^b)_{j\in s}$ and $(e_j^\kappa)_{j\in s}$ and thus of θ. This is also the case for the functions τ and i_0 defined as

$$\begin{aligned}
\tau(\theta) &:= \min_{i\in s} \tau^{i,b}(\theta)\wedge\tau^{i,\kappa}(\theta)\wedge\tau^{i,\partial}(\theta),\\
i_0(\theta) &:= \min\{i \in S,\ \tau^{i,b}(\theta)\wedge\tau^{i,\kappa}(\theta)\wedge\tau^{i,\partial}(\theta) = \tau(\theta)\}.
\end{aligned}$$

For all $\theta \in F$, we define the *final position of* θ, denoted by $Y(\theta) \in (E\cup\partial)^s$, by

$$Y(\theta) := (X_{\tau(\omega)}(\omega_i))_{i\in s},$$

which is also a measurable function of θ.

Then we construct a kernel which, to any given $\theta \in F$ and its final position $Y(\theta)$, associates its position after the branching/killing and selection/resampling event. In order to do so, we define the kernel N^2 from F to $\cup_{s\in\mathcal{P}_f(\mathbb{N})}\{s\}\times(E\cup\partial)^s$ by

$$\begin{aligned}
N^2(\theta) = \mathbf{1}_{\tau(\theta)=\tau^{i_0,b}} p^{b,i_0(\theta)}(Y(\theta))\frac{1}{|s|+1}\sum_{i\in s_+}\delta_{s_+^{-i}}\otimes\delta_{Y(\theta)^{+i_0,-i}}\\
+\mathbf{1}_{\tau(\theta)=\tau^{i_0,b}}\left(1-p^{b,i_0(\theta)}(Y(\theta))\right)\delta_{s_+}\otimes\delta_{Y(\theta)^{+i_0}}\\
+\mathbf{1}_{\tau(\theta)=\tau^{i_0,\kappa}\wedge\tau^{i_0,\partial}}\,p^{\kappa,i_0(\theta)}(Y(\theta))\frac{1}{|s|-1}\sum_{i\in s^{-i_0}}\delta_{s_+^{-i_0}}\otimes\delta_{Y(\theta)^{+i,-i_0}}\\
+\mathbf{1}_{\tau(\theta)=\tau^{i_0,\kappa}\wedge\tau^{i_0,\partial}}\,\left(1-p^{\kappa,i_0(\theta)}(Y(\theta))\right)\delta_{s^{-i_0}}\otimes\delta_{Y(\theta)^{-i_0}}
\end{aligned}$$

where $s_+ = \{1+\max s\}$, $s_+^{-i} = \{1+\max s\}\cup s\setminus\{i\}$, $Y(\theta)^{i_0,-i} \in (E\cup\partial)^{s_+^{-i}}$ is such that

$$Y(\theta)_j^{i_0,-i} = \begin{cases} Y(\theta)_j & \text{for all } j \in s\setminus\{i\},\\ Y(\theta)_{i_0} & \text{for } j = 1+\max s,\end{cases}$$

$Y(\theta)^{i_0} \in (E\cup\partial)^{s_+}$ is such that

$$Y(\theta)_j^{i_0} = \begin{cases} Y(\theta)_j & \text{for all } j \in s,\\ Y(\theta)_{i_0} & \text{for } j = 1+\max s,\end{cases}$$

According to Ionescu Tulcea's extension Theorem, for any initial distribution μ_0 on $F_0 := \cup_{s\in\mathcal{P}_f(\mathbb{N})}\{s\}\times(E\cup\partial)^s$, there exists a probability measure $\mathbb{P}_{\mu_0}$ on $\bar\Omega = F_0\times F^{\mathbb{N}}$, such that, for all $n\geq 2$,

$$\mathbb{P}_{\mu_0}(A_0\in \mathrm{d}a_0,\Theta_1\in\mathrm{d}\theta_1,\,\dots,\,A_{n-1}\in\mathrm{d}a_{n-1},\,\Theta_n\in\mathrm{d}\theta_n)\\ =\mu_0\circ N^1\circ N^2\circ N^1\circ\cdots\circ N^2\circ N^1(\mathrm{d}a_0,\mathrm{d}\theta_1,\cdots,\mathrm{d}a_{n-1},\mathrm{d}\theta_n).$$

Then, we can define, for any $\bar\omega=(a_0,\theta_1,\ldots)$, the sequence of times $(\tau_n)_{n\in\mathbb{N}}$ as

$$\tau_0(\bar\omega)=0 \text{ and } \tau_n(\bar\omega)=\tau(\theta_1)+\cdots+\tau(\theta_n),\ \forall n\geq 0$$

and, for all $t\geq 0$,

$$\bar S_t(\bar\omega)=\sum_{n\geq 0}\mathbf{1}_{\tau_{n-1}\leq t<\tau_n}s(\theta_n).$$

and, for all $i\in\bar S_t(\bar\omega)$,

$$X^i_t(\bar\omega)=X_{t-\tau_{n-1}(\bar\omega)}(\omega_i(\theta_n)).$$

The process $(\bar S_t,(X^i_t)_{i\in\bar S_t})_{t\geq 0}$, defined on the probability space $(\bar\Omega,\mathbb{P})$ and taking its values in F_0, is then the size constrained branching process, which concludes the construction of the process.

Remark 26. Observe that the particle system is well defined without Assumptions 1 and 2. However, if Assumption 1 holds true, then $(\bar S_t,(X^i_t)_{i\in\bar S_t})_{t\geq 0}$ takes its values in $\bigcup_{s\in\mathcal{P}_f(\mathbb{N})}\{s\}\otimes E^s$ almost surely. Also, if Assumption 2 holds true, then $\bar S_t$ is non-empty almost surely, for all $t\in[0,+\infty)$.

The above construction makes use of the fact that X is a measurable adapted process. Using the fact that it is also assumed progressively measurable, we are in position to prove the following.

Proposition 27. *The process $(\bar S_t,(X^i_t)_{i\in\bar S_t})_{t\geq 0}$ is a Markov process with respect to its natural filtration.*

Proof. Fix $t\geq s\geq 0$, and A a measurable subset of F_0. For all $k\geq 1$, all $s_1<s_2<\cdots<s_k=s$ and all measurable subsets $A_1,\ldots,A_k$ of F_0, we have, setting $\mathbb{X}_t := (\bar S_t,(X^i_t)_{i\in\bar S_t})$,

$$\mathbb{P}_{\mu_0}\left(\mathbb{X}_{s_1}\in A_1,\ldots,\mathbb{X}_{s_k}\in A_k,\mathbb{X}_t\in A\right)\\ =\sum_{n_1\leq\cdots\leq n_k=m\leq n}\mathbb{P}_{\mu_0}\left(\mathbb{X}_{s_1}\in A_1,\ldots,\mathbb{X}_{s_k}\in A_k,\ s_i\in[\tau_{n_i-1},\tau_{n_i}),\ \mathbb{X}_t\in A,\ t\in[\tau_n,\tau_{n+1})\right). \quad (60)$$

Then, conditioning with respect to $\mathcal{G}_m:=\sigma(\Theta_1,\ldots,\Theta_m,A_m)$, we have

$$\mathbb{P}_{\mu_0}\left(\mathbb{X}_{s_1}\in A_1,\ldots,\mathbb{X}_{s_k}\in A_k,\ s_i\in[\tau_{n_i-1},\tau_{n_i}),\ \mathbb{X}_{s+t}\in A,\ s+t\in[\tau_n,\tau_{n+1})\mid\mathcal{G}_m\right)\\ =\mathbf{1}_{\mathbb{X}_{s_1}\in A_1,\ldots,\mathbb{X}_{s_\ell}\in A_\ell,\,s_i\in[\tau_{n_i-1},\tau_{n_i}),\tau_m\leq s}\times f(A_m,\tau_m), \quad (61)$$

where

$$f(a,u):=\mathbb{P}_a\left(\mathbb{X}_{s_{\ell+1}-u}\in A_{\ell+1},\ldots,\mathbb{X}_{s_k-u}\in A_k,\ s_i-u\in[0,\tau_1),\mathbb{X}_{t-u}\in A,\ t-u\in[\tau_{n-m},\tau_{n-m+1})\right).$$

There are now two possible situations, either $m=n$ or $m<n$, which we consider separately.

Case $m = n$. In the former case, we observe that, by definition of N^1,

$$
\begin{aligned}
f(a,u) &= \mathbb{P}_a\left(\mathbb{X}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{X}_{s_k-u}\in A_k,\ s_i-u\in[0,\tau_1),\mathbb{X}_{t-u}\in A,\ t-u\in[0,\tau_1)\right)\\
&= \mathtt{P}_a\left(\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k,\ s_i-u\in[0,\mathtt{t}),\mathbb{Y}_{t-u}\in A,\ t-u\in[0,\mathtt{t})\right),
\end{aligned}
$$

where $\mathtt{P}_a$ is the product measure $\Omega' = \mathbb{P}^{\otimes s}\otimes\mathcal{E}(1)^s$ on the probability space $\Omega^s(\mathbb{R}_+)^s\otimes(\mathbb{R}_+)^s$ endowed with the filtration defined by $\mathcal{H}_t := \mathcal{F}_t^{\otimes s}\otimes\{\emptyset,\mathbb{R}_+\}^s\otimes\{\emptyset,\mathbb{R}_+\}^s$, with the variables $\mathbb{Y}$ and $\mathtt{t}$ being defined, for all $\theta' = ((\omega_i)_{i\in s},(e_i^b)_{i\in s},(e_i^\kappa)_{i\in s})\in\Omega'$, by

$$
\mathbb{Y}_t = (s,(X_t(\omega_i)_{i\in s}),\ \forall t\geq 0,
$$

and $\mathtt{t} = \min_{i\in s}\mathtt{t}^{b,i}\wedge\mathtt{t}^{\kappa,i}\wedge t^{\partial,i}$, with

$$
\begin{aligned}
\mathtt{t}^{b,i} &= \inf\{t\geq 0,\ \int_0^t b^{i_0}((X_u(\omega_j))_{j\in s})\,\mathrm{d}u = e_i^b\},\\
\mathtt{t}^{\kappa,i} &= \inf\{t\geq 0,\ \int_0^t \kappa^{i_0}((X_u(\omega_j))_{j\in s})\,\mathrm{d}u = e_i^\kappa\},\\
\mathtt{t}^{\partial,i} &= \inf\{t\geq 0,\ X_t(\omega_i)\in\partial\}.
\end{aligned}
$$

We emphasise that the second component of $\mathbb{Y}$ is distributed as a vector indexed by s of independent copies of X starting from x_i, for all $i\in s$, and that it is a progressively measurable Markov process with respect to the filtration $\mathcal{H}$. In addition, the random variables e_i^b and e_i^κ, $i\in s$, are independent from each others and from $\mathcal{H}_t$, for all $t\geq 0$.

Using the Markov property at time $s-u$ for the process $\mathbb{Y}$, we obtain

$$
\begin{aligned}
&\mathtt{P}_a\left(\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k,\ s_i-u\in[0,\mathtt{t}),\mathbb{Y}_{t-u}\in A,\ t-u\in[0,\mathtt{t})\right)\\
&\quad= \mathtt{E}_a\left(\mathbf{1}_{\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k}\mathbf{1}_{\mathbb{Y}_{t-u}\in A\setminus\partial}\,e^{-\int_0^{t-u}h(\mathbb{Y}_v)\,\mathrm{d}v}\right)\\
&\quad= \mathtt{E}_a\left(\mathbf{1}_{\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k}\mathtt{E}_a\left(\mathbf{1}_{\mathbb{Y}_{t-u}\in A\setminus\partial}\,e^{-\int_0^{t-u}h(\mathbb{Y}_v)\,\mathrm{d}v}\mid\mathcal{H}_{s-u}\right)\right)\\
&\quad= \mathtt{E}_a\left(\mathbf{1}_{\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k}e^{-\int_0^{s-u}h(\mathbb{Y}_v)\,\mathrm{d}v}\mathtt{E}_{\mathbb{Y}_{s-u}}\left(\mathbf{1}_{\mathbb{Y}_{t-s}\in A\setminus\partial}\,e^{-\int_0^{t-s}h(\mathbb{Y}_v)\,\mathrm{d}v}\right)\right)\\
&\quad= \mathtt{E}_a\left(\mathbf{1}_{\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k}\mathbf{1}_{s_k-u\in[0,\mathtt{t})}\mathtt{E}_{\mathbb{Y}_{s-u}}\left(\mathbf{1}_{\mathbb{Y}_{t-s}\in A\setminus\partial}\,\mathbf{1}_{t-s\in[0,\mathtt{t})}\right)\right)
\end{aligned}
$$

where $h(s,(x_i)_{i\in s}) := \sum_{i\in s}\kappa^i(x_i)+b^i(x_i)$, and where we used the fact that $\mathbb{Y}$ is progressively measurable to deduce that $\int_0^{s-u}h(\mathbb{Y}_v)\,\mathrm{d}v$ is $\mathcal{H}_{s-u}$-measurable. We deduce that

$$
f(a,u) = \mathbb{E}_a\left(\mathbf{1}_{\mathbb{X}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{X}_{s_k-u}\in A_k,\,s_i-u\in[0,\tau_1)}\,\mathbb{E}_{\mathbb{X}_{s-u}}\left(\mathbb{X}_{t-s}\in A,\ t-s\in[0,\tau_1)\right)\right)
$$

Taking the expectation in (61), we finally deduce that

$$
\begin{aligned}
&\mathbb{P}_{\mu_0}\left(\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_k}\in A_k,\ s_i\in[\tau_{n_i-1},\tau_{n_i}),\ \mathbb{X}_{s+t}\in A,\ s+t\in[\tau_n,\tau_{n+1})\right)\\
&\qquad\qquad= \mathbb{E}_{\mu_0}\left(\mathbf{1}_{\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_\ell}\in A_\ell,\,s_i\in[\tau_{n_i-1},\tau_{n_i}),\tau_m\leq s}\times f(A_m,\tau_m)\right),
\end{aligned}
$$

and then, by definition of the law of $(A_0,\Theta_1,\cdots,A_{m-1},\Theta_m,A_m,\dots)$, we deduce that

$$
\begin{aligned}
&\mathbb{P}_{\mu_0}\left(\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_k}\in A_k,\ s_i\in[\tau_{n_i-1},\tau_{n_i}),\ \mathbb{X}_{s+t}\in A,\ s+t\in[\tau_n,\tau_{n+1})\right)\\
&\qquad= \mathbb{E}_{\mu_0}\left(\mathbf{1}_{\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_k}\in A_k,\,s_i\in[\tau_{n_i-1},\tau_{n_i})}\mathbb{P}_{\mathbb{X}_s}\left(\mathbb{X}_{t-s}\in A,\ t-s\in[0,\tau_1)\right)\right). \quad (62)
\end{aligned}
$$

Case $m < n$. We can now proceed with the case $m < n$. In this case, we write

$$f(a,u) = \mathbb{E}_a\left(\mathbf{1}_{\mathbb{X}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{X}_{s_k-u}\in A_k,\, s_i-u\in[0,\tau_1)}\, g_u(A_1,\tau_1)\right),$$

where

$$g_u(b,v) := \mathbb{E}_{A_1}\left(\mathbb{X}_{t-u-v}\in A,\, t-u-v\in[\tau_{n-m-1},\tau_{n-m})\right)$$

Using the same strategy as above (with the same definition of $\mathbb{Y}$), we obtain, using also the Markov property at time $s_k - u = s - u$,

$$\begin{aligned} f(a,u) &= \mathtt{E}_a\left(\mathbf{1}_{\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k,\, s_i-u\in[0,\mathtt{t})}\, g_u(\mathbb{Y}_{\mathtt{t}},\mathtt{t})\right)\\ &= \mathtt{E}_a\left(\mathbf{1}_{\mathbb{Y}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{Y}_{s_k-u}\in A_k,\, s_i-u\in[0,\mathtt{t})}\, \mathtt{E}_{\mathbb{Y}_{s-u}}(g_u(\mathbb{Y}_{\mathtt{t}},\mathtt{t}))\right)\\ &= \mathbb{E}_a\left(\mathbf{1}_{\mathbb{X}_{s_{\ell+1}-u}\in A_{\ell+1},\dots,\mathbb{X}_{s_k-u}\in A_k,\, s_i-u\in[0,\tau_1)}\, \mathbb{E}_{\mathbb{X}_{s-u}}(g_u(\mathbb{X}_{\tau_1},\tau_1))\right) \end{aligned}$$

Using again the definition of the law of $(A_0,\Theta_1,\cdots,A_{m-1},\Theta_m,A_m,\dots)$, we deduce that

$$\begin{aligned} &\mathbb{P}_{\mu_0}\left(\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_k}\in A_k,\, s_i\in[\tau_{n_i-1},\tau_{n_i}),\, \mathbb{X}_{s+t}\in A,\, s+t\in[\tau_n,\tau_{n+1})\right)\\ &\qquad = \mathbb{E}_{\mu_0}\left(\mathbf{1}_{\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_k}\in A_k,\, s_i\in[\tau_{n_i-1},\tau_{n_i})}\mathbb{P}_{\mathbb{X}_s}\left(\mathbb{X}_{t-s}\in A,\, t-s\in[\tau_{n-m},\tau_{n-m+1})\right)\right). \quad (63) \end{aligned}$$

Conclusion. Finally, using (62) and (63) and summing over the indices n_i and n as in (60), we deduce that

$$\mathbb{P}_{\mu_0}\left(\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_k}\in A_k,\mathbb{X}_t\in A\right) = \mathbb{E}_{\mu_0}\left(\mathbf{1}_{\mathbb{X}_{s_1}\in A_1,\dots,\mathbb{X}_{s_k}\in A_k}\,\mathbb{P}_{\mathbb{X}_s}\left(\mathbb{X}_{t-s}\in A\right)\right).$$

Since the property holds true for all finite family of time indices $s_1 < \ldots < s_k = s$ and all $A_1,\dots,A_k$, we deduce that

$$\mathbb{P}_{\mu_0}\left(\mathbb{X}_t\in A \mid \sigma(\mathbb{X}_u,\, u\le s)\right) = \mathbb{P}_{\mathbb{X}_s}\left(\mathbb{X}_{t-s}\in A\right),$$

which concludes the proof of the proposition. □

Remark 28. Although we won't use this property, the process $(\bar{S}_t,(X^i_t)_{t\in\bar{S}_t})_{t\ge 0}$ can be shown to be a Markov process with respect to larger filtrations. On a related note, it is natural to ask whether the strong Markov property for X entails the strong Markov property for the particle system. This is not clear at all in full generality and we leave it as an open problem. However, it can be shown, using a similar approach as in [39], that, under additional regularity properties on the process X and on the functionals b,κ, and additional requirements on the state space E, the particle system is indeed a strong Markov process in a quite general setting.

References

[1] V. Bansaye. Spine for interacting populations and sampling. arXiv preprint arXiv:2105.03185, 2021.

[2] M. Barczy, D. Bezdány, and G. Pap. A note on asymptotic behavior of critical Galton-Watson processes with immigration. arXiv e-prints, page arXiv:2103.07878, Mar. 2021.

[3] J. Bertoin. On a Feynman-Kac approach to growth-fragmentation semigroups and their asymptotic behaviors. Journal of Functional Analysis, 277(11):108270, 2019.

[4] J. Bertoin et al. Markovian growth-fragmentation processes. Bernoulli, 23(2):1082–1101, 2017.

[5] M. Bieniek, K. Burdzy, and S. Pal. Extinction of Fleming-Viot-type particle systems with strong drift. Electron. J. Probab., 17:no. 11, 15, 2012.

[6] K. Burdzy, R. Holyst, D. Ingerman, and P. March. Configurational transition in a Fleming-Viot-type model and probabilistic interpretation of Laplacian eigenfunctions. J. Phys. A, 29(29):2633–2642, 1996.

[7] K. Burdzy, R. Hołyst, and P. March. A Fleming-Viot particle representation of the Dirichlet Laplacian. Comm. Math. Phys., 214(3):679–703, 2000.

[8] F. Cérou, B. Delyon, A. Guyader, and M. Rousset. On the Asymptotic Normality of Adaptive Multilevel Splitting. SIAM/ASA Journal on Uncertainty Quantification, 7(1):1–30, 2019. 38 pages, 5 figures.

[9] N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais. Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions. In Séminaire de Probabilités XLIX, pages 165–182. Springer International Publishing, Cham, 2018.

[10] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. Probab. Theory Related Fields, 164(1):243–283, 2016.

[11] N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed markov processes. Stochastic Processes and their Applications, 135:51–74, 2021.

[12] B. Cloez and J. Corujo. Uniform in time propagation of chaos for a Moran model. arXiv e-prints, page arXiv:2107.10794, July 2021.

[13] B. Cloez and M.-N. Thai. Quantitative results for the Fleming-Viot particle system and quasi-stationary distributions in discrete space. Stochastic Process. Appl., 126(3):680–702, 2016.

[14] A. M. Cox, S. C. Harris, E. L. Horton, and A. E. Kyprianou. Multi-species neutron transport equation. Journal of Statistical Physics, 176(2):425–455, 2019.

[15] A. M. Cox, S. C. Harris, A. E. Kyprianou, and M. Wang. Monte-Carlo methods for the neutron transport equation. arXiv preprint arXiv:2012.02864, 2020.

[16] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. J. Appl. Probab., 4:192–196, 1967.

[17] D. A. Dawson. Measure-valued Markov processes. In École d'Été de Probabilités de Saint-Flour XXI—1991, volume 1541 of Lecture Notes in Math., pages 1–260. Springer, Berlin, 1993.

[18] P. Del Moral. Feynman-Kac formulae. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications.

[19] P. Del Moral. Mean field simulation for Monte Carlo integration, volume 126 of Monographs on Statistics and Applied Probability. CRC Press, Boca Raton, FL, 2013.

[20] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of feynman-kac formulae with applications to non-linear filtering. pages 1–145, 2000.

[21] P. Del Moral and L. Miclo. A Moran particle system approximation of Feynman–Kac formulae. Stochastic processes and their applications, 86(2):193–216, 2000.

[22] P. Del Moral and L. Miclo. Asymptotic results for genetic algorithms with applications to nonlinear estimation. pages 439–493, 2001.

[23] P. Del Moral and L. Miclo. On the stability of nonlinear Feynman-Kac semigroups. Ann. Fac. Sci. Toulouse Math. (6), 11(2):135–175, 2002.

[24] P. Del Moral and L. Miclo. Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman-Kac semigroups. ESAIM. Probability and Statistics, 7:171–208, Mar. 2003.

[25] P. Del Moral and D. Villemonais. Exponential mixing properties for time inhomogeneous diffusion processes with killing. ArXiv e-prints, Dec. 2015.

[26] P. Del Moral and D. Villemonais. Exponential mixing properties for time inhomogeneous diffusion processes with killing. Bernoulli, 24(2):1010–1032, 2018.

[27] B. Delyon, F. Cérou, A. Guyader, and M. Rousset. A central limit theorem for Fleming-Viot particle systems with hard killing. arXiv preprint arXiv:1709.06771, 2017.

[28] A. Etheridge. Some Mathematical Models from Population Genetics: École D'Été de Probabilités de Saint-Flour XXXIX-2009, volume 2012. Springer Science & Business Media, 2011.

[29] P. A. Ferrari and N. Marić. Quasi stationary distributions and Fleming-Viot processes in countable spaces. Electron. J. Probab., 12:no. 24, 684–702 (electronic), 2007.

[30] I. Garcia, G. Yanev, M. Molina, N. Yanev, and M. Velasco. Continuous-time controlled branching processes. Comptes rendus de l'Académie bulgare des Sciences, 2021.

[31] M. González, M. Molina, and I. D. Puerto. On the class of controlled branching processes with random control functions. Journal of Applied Probability, 39(4):804–815, 2002.

[32] I. Grigorescu and M. Kang. Hydrodynamic limit for a Fleming-Viot type system. Stochastic Process. Appl., 110(1):111–143, 2004.

[33] I. Grigorescu and M. Kang. Ergodic properties of multidimensional Brownian motion with rebirth. Electron. J. Probab., 12:no. 48, 1299–1322, 2007.

[34] P. Groisman and M. Jonckheere. Simulation of quasi-stationary distributions on countable spaces. Markov Process. Related Fields, 19(3):521–542, 2013.

[35] S. C. Harris and M. I. Roberts. The many-to-few lemma and multiple spines. In Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, volume 53, pages 226–242. Institut Henri Poincaré, 2017.

[36] E. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic methods for the neutron transport equation I: Linear semigroup asymptotics. six figures, Oct. 2018.

[37] N. Ikeda, M. Nagasawa, S. Watanabe, et al. Branching Markov processes I. Journal of Mathematics of Kyoto University, 8(2):233–278, 1968.

[38] N. Ikeda, M. Nagasawa, S. Watanabe, et al. Branching Markov processes II. Journal of Mathematics of Kyoto University, 8(3):365–410, 1968.

[39] N. Ikeda, M. Nagasawa, S. Watanabe, et al. Branching Markov processes III. Journal of Mathematics of Kyoto University, 9(1):95–160, 1969.

[40] P. Jagers. General branching processes as Markov fields. Stochastic Processes and their Applications, 32(2):183–212, 1989.

[41] P. Jagers. Branching processes as population dynamics. Bernoulli, 1(1-2):191–200, 1995.

[42] K. Kawazu and S. Watanabe. Branching processes with immigration and related limit theorems. Teor. Verojatnost. i Primenen., 16:34–51, 1971.

[43] A. Lambert. The branching process with logistic growth. Ann. Appl. Probab., 15(2):1506–1535, 2005.

[44] A. Lambert. Population dynamics and random genealogies. Stochastic Models, 24(sup1):45–163, 2008.

[45] T. Lelièvre, L. Pillaud-Vivien, and J. Reygner. Central limit theorem for stationary Fleming-Viot particle systems in finite spaces. ALEA, Lat. Am. J. Probab. Math. Stat., 15(2):1163–1182, 2018.

[46] Z.-h. Li. Branching processes with immigration and related topics. Frontiers of Mathematics in China, 1(1):73–97, 2006.

[47] J. M. Marshall. A branching process for the early spread of a transposable element in a diploid population. Journal of mathematical biology, 57(6):811, 2008.

[48] P. A. P. Moran. Random processes in genetics. In Mathematical proceedings of the cambridge philosophical society, volume 54, pages 60–71. Cambridge University Press, 1958.

[49] W. Oçafrain and D. Villemonais. Convergence of a non-failable mean-field particle system. Stoch. Anal. Appl., 35(4):587–603, 2017.

[50] P. Olofsson. General branching processes with immigration. Journal of Applied Probability, 33(4):940–948, 1996.

[51] M. Rousset. On the control of an interacting particle estimation of Schrödinger ground states. SIAM J. Math. Anal., 38(3):824–844 (electronic), 2006.

[52] B. A. Sevastyanov and A. M. Zubkov. Controlled branching processes. Theory of Probability & Its Applications, 19(1):14–24, 1974.

[53] O. Tough and J. Nolen. The Fleming-Viot process with McKean-Vlasov dynamics, 2020.

[54] V. A. Vatutin. A critical Galton–Watson branching process with emigration. Theory of Probability & Its Applications, 22(3):465–481, 1978.

[55] M. G. Velasco, I. M. D. P. García, and G. P. Yanev. Controlled branching processes. John Wiley & Sons, 2017.

[56] D. Villemonais. Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift. Electronic Journal of Probability, 16:1663–1692, 2011.

[57] D. Villemonais. General approximation method for the distribution of Markov processes conditioned not to be killed. ESAIM Probab. Stat., 18:441–467, 2014.

[58] G. P. Yanev. Critical controlled branching processes and their relatives. arXiv preprint arXiv:1411.6045, 2014.

[59] N. M. Yanev. Conditions for degeneracy of φ-branching processes with random φ. Theory of Probability & Its Applications, 20(2):421–428, 1976.

[60] N. M. Yanev. Branching processes in cell proliferation kinetics. In Workshop on Branching Processes and Their Applications, pages 159–178. Springer, 2010.