Monte Carlo Methods for the Neutron Transport Equation[*]

Alexander M. G. Cox[†], Simon C. Harris[‡], Andreas E. Kyprianou[†], and Minmin Wang[§]

Abstract. This paper continues our treatment of the neutron transport equation (NTE), building on the work in [A. M. G. Cox et al., *J. Stat. Phys.*, 176 (2019), pp. 425–455; E. Horton, A. E. Kyprianou, and D. Villemonais, *Ann Appl. Probab.*, 30 (2020), pp. 2573–2612; and S. C. Harris, E. Horton, and A. E. Kyprianou, *Ann. Appl. Probab.*, 30 (2020), pp. 2815–2845], which describes the density (equivalently, flux) of neutrons through inhomogeneous fissile media. Our aim is to analyze existing and novel Monte Carlo (MC) algorithms, aimed at simulating the lead eigenvalue associated with the underlying model. This quantity is of principal importance in the nuclear regulatory industry, for which the NTE must be solved on complicated inhomogeneous domains corresponding to nuclear reactor cores, irradiative hospital equipment, food irradiation equipment, and so on. We include a complexity analysis of such MC algorithms, noting that no such undertaking has previously appeared in the literature. The new MC algorithms offer a variety of advantages and disadvantages of accuracy versus cost, as well as the possibility of more convenient computational parallelization.

Key words. neutron transport equation, principal eigenvalue, semigroup theory, Perron–Frobenius decomposition, Monte Carlo simulation, complexity, Doob h-transform, twisted Monte Carlo

AMS subject classifications. 82D75, 60J80, 60J75, 60J99

DOI. 10.1137/21M1390578

1. Introduction: The neutron transport equation. In this paper, we continue the stochastic analysis of the neutron transport equation (NTE); cf. [29, 26, 13], though here we focus mainly on Monte Carlo methods. We begin in this first section with a review of the classical and recent contributions to the mathematical framework of the NTE, from which we will build completely new Monte Carlo algorithms in the subsequent sections.

We recall that the NTE describes the density (or equivalently flux[1]) of neutrons traveling in an inhomogeneous fissile medium where nuclear fission is actively creating additional neutrons. Neutron density is described as a function of time, t, Euclidean location, r, neutron velocity, and neutron energy, E. In the physics literature, it is not unusual to assume, as indeed we shall do here, that energy is a function of velocity, thereby reducing the number of variables by one. This allows us to describe the dependency of neutron density more simply in terms

[†]Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK (a.m.g.cox@bath.ac.uk a.kyprianou@bath.ac.uk).

[‡]Department of Statistics, University of Auckland, Auckland 1142, New Zealand (simon.harris@auckland.ac.nz).

[§]School of Mathematical and Physical Sciences, University of Sussex, Brighton, BN1 9RH, UK (wangminmin03@gmail.com).

[1]flux = speed × density

of time and of what we call the *configuration variables* $(r, v) \in D \times V$, where $D \subseteq \mathbb{R}^3$ is a nonempty, open, smooth, bounded, and convex domain, and V is the velocity space, which we will assume to be $V = \{v \in \mathbb{R}^3 : v_{\min} \leq |v| \leq v_{\max}\}$, where $0 < v_{\min} < v_{\max} < \infty$.

While the NTE is usually stated in the applied mathematics and physics literature in its forwards form, we will introduce it here in its backwards form. A fuller description of the relationship between the two is given in the accompanying papers [13] and [29]. From a mathematical perspective, the backwards NTE for neutron density is given[2] by

$$
\frac{\partial}{\partial t}\psi_t(r, v) = v \cdot \nabla \psi_t(r, v) - \sigma(r, v)\psi_t(r, v)
$$

$$
(1.1) \qquad + \sigma_{\mathtt{s}}(r, v) \int_V \psi_t(r, v')\pi_{\mathtt{s}}(r, v, v')\mathrm{d}v' + \sigma_{\mathtt{f}}(r, v) \int_V \psi_t(r, v')\pi_{\mathtt{f}}(r, v, v')\mathrm{d}v',
$$

where the different components (or *cross-sections*[3] as they are known in the nuclear physics literature) have the following interpretation:

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

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

$\sigma(r, v)$: the sum of the rates $\sigma_{\mathtt{f}} + \sigma_{\mathtt{s}}$, also known as the total cross-section;

$\pi_{\mathtt{s}}(r, v, v')\mathrm{d}v'$: the scattering yield at velocity v' from incoming velocity v,

satisfying $\int_V \pi_{\mathtt{s}}(r, v, v')\mathrm{d}v' = 1$; and

$\pi_{\mathtt{f}}(r, v, v')\mathrm{d}v'$: the neutron yield at velocity v' from fission with incoming velocity v,

satisfying $m_{\mathtt{f}}(r, v) := \int_V \pi_{\mathtt{f}}(r, v, v')\mathrm{d}v' < \infty$.

Note that scattering is the physical process that occurs when a neutron comes into close proximity with an atomic nucleus, causing a change in velocity. A justification of the structure of the backwards (and indeed the forwards) NTE form from the physics of nuclear scattering and fission is given in a number of classical texts; see, e.g., [15] and [41], and also [13] for a multispecies treatment.

It is also usual to assume the additional boundary conditions

$$
(1.2) \qquad \begin{cases} \psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \\ \psi_t(r, v) = 0 & \text{for } r \in \partial D \text{ if } v \cdot \mathbf{n}_r > 0, \end{cases}
$$

---

[2]The differential operator $\nabla \psi_t(r, v)$ is the gradient in the spatial variable, r, only, so that $v \cdot \nabla \psi_t(r, v)$ is the gradient of ψ in the direction of v, multiplied by the norm of the velocity.

[3]There is a discrepancy between what we mean by cross-sections here and what is typical in the physics and engineering literature. Here, cross-sections $\sigma_{\mathtt{s}}, \sigma_{\mathtt{f}}$ are considered as rates per unit time, which is natural for probabilistic considerations, whereas the norm is to consider them as rates per unit neutron track length, which is natural when thinking of modeling how physical interaction occurs. The difference is a factor of v, which converts rate per unit track length to rate per unit time. This also explains why (1.1) is usually written with a factor $1/v$ preceding the time derivative.

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 which we will later assume has some additional properties. Roughly speaking, this means that neutrons at the boundary which are traveling in the direction of the exterior of the domain are lost to the system. The second of the two conditions in (1.2) is often written $\psi_t|_{\partial D^+} = 0$, where $\partial D^+ := \{(r, v) \in D \times V : r \in \partial D \text{ if } v \cdot \mathbf{n}_r > 0\}$.

For mathematical reasons, in the forthcoming analysis, we will assume the following:

(H1): Cross-sections $\sigma_{\mathtt{s}}$, $\sigma_{\mathtt{f}}$, $\pi_{\mathtt{s}}$, and $\pi_{\mathtt{f}}$ are uniformly bounded away from infinity.

(H2): The mixed cross-sections satisfy

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

(H3): There exists an open ball B, compactly embedded in D, such that

$$\inf_{r \in B, v, v' \in V} \sigma_{\mathtt{f}}(r, v)\pi_{\mathtt{f}}(r, v, v') > 0.$$

Roughly speaking, (H2) says that either scattering or fission, or both, must be able to occur everywhere at all energies. On the other hand, (H3) says that fission must be able to occur at least in some region at all energies.

In addition, the maximum number of neutrons that can be emitted during a fission event with positive probability is capped by physical constraints. For example, in an environment where the heaviest nucleus is *uranium-235*, there are at most 143 neutrons that can be released in a fission event, although in reality it is more likely that 2 or 3 are released in any fission event. We will thus occasionally work with the following assumption:

(H4): Fission offspring are bounded in number by the constant $n_{\mathtt{max}} > 1$.

In particular this means that $\sup_{r \in D, v \in V} m_{\mathtt{f}}(r, v) \leq n_{\mathtt{max}}$.

It turns out that (1.1) is ill-defined unless considered in an appropriate sense. There are in fact two ways this can be done. The first approach is to formulate the problem as an abstract Cauchy problem (ACP) in an appropriate Banach space; see, e.g., [14, 15], which is typically the approach taken in the applied mathematics literature. The second approach is to consider the mild form of the equation. This approach has been paid somewhat less attention, although it lends itself better to Monte Carlo methods.

2. Mild NTE and stochastic representation. In order for us to give a clear statement of what a mild version of (1.1) should look like and how it relates to Monte Carlo simulation, we must first consider an auxiliary stochastic process.

2.1. The physical process. Consider a neutron branching process (NBP), which at time $t \geq 0$ is represented by a configuration of particles which are specified via their physical location and velocity in $D \times V$, say $\{(r_i(t), v_i(t)) : i = 1, \dots, N_t\}$, where N_t is the number of particles alive at time $t \geq 0$. In order to describe the process, we will represent it as a process in the space of the random counting measures

$$X_t(A) = \sum_{i=1}^{N_t} \delta_{(r_i(t), v_i(t))}(A), \qquad A \in \mathcal{B}(D \times V),\ t \geq 0,$$

where $\delta_{(r,v)}$ is the Dirac measure of unit mass at (r,v) defined on $\mathcal{B}(D \times V)$, the Borel subsets of $D \times V$. The evolution of $(X_t, t \geq 0)$ is a stochastic process valued in the space of atomic measures $\mathcal{M}(D \times V) := \{\sum_{i=1}^{n} \delta_{(r_i, v_i)} : n \in \mathbb{N}, (r_i, v_i) \in D \times V, i = 1, \ldots, n\}$ which evolves randomly as follows.

A particle positioned at r with velocity v will continue to move along the trajectory $r + vt$ for $t \geq 0$ until one of the following things happens:

(i) The particle leaves the physical domain D, in which case it is instantaneously killed.

(ii) Independently of all other neutrons, a scattering event occurs when a neutron comes in close proximity to an atomic nucleus and, accordingly, makes an instantaneous change of velocity. For a neutron in the system with position and velocity (r,v), if we write $T_{\mathtt{s}}$ for the random time in which scattering may occur, then independently of any other physical event that may affect the neutron, $\Pr(T_{\mathtt{s}} > t) = \exp\{-\int_0^t \sigma_{\mathtt{s}}(r + vs, v)\mathrm{d}s\}$ for $t \geq 0$ with $(r + vs : s \in [0,t]) \subset D$, that is, as long as the particle has remained within the domain up to time t.

When scattering occurs at space-velocity (r, v), the new velocity v' is selected in V independently with probability $\pi_{\mathtt{s}}(r, v, v')\mathrm{d}v'$.

(iii) Independently of all other neutrons, a fission event occurs when a neutron smashes into an atomic nucleus. For a neutron in the system with initial position and velocity (r,v), if we write $T_{\mathtt{f}}$ for the random time that fission may occur, then independently of any other physical event that may affect the neutron, $\Pr(T_{\mathtt{f}} > t) = \exp\{-\int_0^t \sigma_{\mathtt{f}}(r + vs, v)\mathrm{d}s\}$ for $t \geq 0$ with $(r + vs : s \in [0,t]) \subset D$.

When fission occurs, the smashing of the atomic nucleus produces lower mass isotopes and releases a random number of neutrons, say $N \geq 0$, which are ejected from the point of impact with randomly distributed, and possibly correlated, velocities, say $\{v_i : i = 1, \ldots, N\}$. The outgoing velocities are described by the atomic random measure

$$(2.1) \qquad \mathcal{Z}(A) := \sum_{i=1}^{N} \delta_{v_i}(A), \qquad A \in \mathcal{B}(V).$$

When fission occurs at location $r \in \mathbb{R}^d$ from a particle with incoming velocity $v \in V$, we denote by $\mathcal{P}_{(r,v)}$ the law of $\mathcal{Z}$. The probability laws $\mathcal{P}_{(r,v)}$, with their corresponding expectations $\mathcal{E}_{(r,v)}$, are such that, for $v' \in V$ and bounded measurable $g : V \to [0, \infty)$,

$$(2.2) \qquad \int_V g(v')\pi_{\mathtt{f}}(r, v, v')\mathrm{d}v' = \mathcal{E}_{(r,v)}\left[\int_V g(v')\mathcal{Z}(\mathrm{d}v')\right] =: \mathcal{E}_{(r,v)}[\langle g, \mathcal{Z}\rangle].$$

Note the possibility that $\Pr(N = 0) > 0$, which will be tantamount to neutron capture (that is, where a neutron slams into a nucleus but no fission results and the neutron is absorbed into the nucleus).

In essence, one may think of the process $X := (X_t, t \geq 0)$ as a typed spatial Markov branching process, where the type of each particle is the velocity $v \in V$ and the underlying motion is nothing more than movement in a straight line with velocity v until the next jump in type (scattering) or branching event (fission), with these occurring independently for each particle at rates depending only on their current spatial position and type.

Remark 2.1. The NBP is thus parameterized by the quantities $\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}$ and the family of measures $\mathcal{P} = (\mathcal{P}_{(r,v)}, r \in D, v \in V)$, and accordingly we refer to it as a $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \mathcal{P})$-NBP. It is associated to the NTE via the relation (2.2); however, this association does not uniquely identify the NBP. Said another way, the quantities $\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}}$ in the NTE do not uniquely identify an underlying NBP. In particular, only the average number of particles created at fission events is specified via $\pi_{\mathtt{f}}$, but not their probability distributions as required for $\mathcal{P}$. See Remark 2.1 of [29].

Later on, we will see that it is possible to simulate an NBP with only the parameters $\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}}$ given. Such processes are not uniquely determined; however, the aforementioned cross-sections are sufficient to identify at least one NBP which is associated to the NTE with those cross-sections. In that case, we will abuse our terminology and refer to it as a $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}})$-NBP.

Write $\mathbb{P}_{\delta_{(r,v)}}$ for the the law of X when issued from a single particle with space-velocity configuration $(r, v) \in D \times V$. More generally, for $\mu \in \mathcal{M}(D \times V)$, we understand $\mathbb{P}_\mu := \mathbb{P}_{\delta_{(r_1,v_1)}} \otimes \cdots \otimes \mathbb{P}_{\delta_{(r_n,v_n)}}$ when $\mu = \sum_{i=1}^n \delta_{(r_i,v_i)}$. In other words, the process X when issued from initial configuration μ is equivalent to issuing n independent copies of X, where the ith copy starts from a single particle with configuration (r_i, v_i) for $i = 1, \ldots, n$.

Like all spatial Markov branching processes, $(X, \mathbb{P})$, where $\mathbb{P} := (\mathbb{P}_\mu, \mu \in \mathcal{M}(D \times V))$, respects the *Markov branching property* with respect to the filtration $\mathcal{F}_t := \sigma((r_i(s), v_i(s)) : i = 1, \ldots, N_s, s \leq t), t \geq 0$. That is to say, for all bounded and measurable $g : D \times V \to [0, \infty)$ and $\mu \in \mathcal{M}(D \times V)$ written $\mu = \sum_{i=1}^n \delta_{(r_i,v_i)}$ for $s, t \geq 0$, on the event $\{X_t = \mu\} \in \mathcal{F}_t$, we have $\mathbb{E}_\mu[\prod_{i=1}^{N_{t+s}} g(r_i(t+s), v_i(t+s))|\mathcal{F}_t] = \prod_{i=1}^n u_s[g](r_i, v_i)$ for $t \geq 0, r_i \in D, v_i \in V$, where $u_s[g](r,v) := \mathbb{E}_{\delta_{(r,v)}}[\prod_{i=1}^{N_s} g(r_i(s), v_i(s))]$.

What is of particular interest to us in the context of the NTE is what we call the *expectation semigroup* of the NBP. More precisely, and with preemptive notation, we are interested in

$$(2.3) \qquad \psi_t[g](r,v) := \mathbb{E}_{\delta_{(r,v)}}[\langle g, X_t \rangle], \qquad t \geq 0, r \in \bar{D}, v \in V,$$

for $g \in L_\infty^+(D \times V)$, the space of nonnegative uniformly bounded measurable functions on $D \times V$. Here we have made a slight abuse of notation (see $\langle \cdot, \cdot \rangle$ as it appears in (2.2)) and written $\langle g, X_t \rangle$ to mean $\int_{D \times V} g(r, v) X_t(\mathrm{d}r, \mathrm{d}v)$, that is, $\langle g, X_t \rangle = \sum_{i=1}^{N_t} g(r_i(t), v_i(t))$.

Thus, $\psi_t[g](r,v)$ gives the average value of a weighted sum over all the neutrons present at time t when started from a single neutron with a configuration (r, v), where the weight of a neutron with configuration (r', v') is given by $g(r', v')$.

Heuristically, if we take $g(r', v') = \delta_{(r',v')}$, then ψ_t recovers the expected density of neutrons that have a position-velocity configuration (r', v') at time t. In particular, the expected flux of neutrons across any planar cross-section could then be recovered from these quantities.

To see why $(\psi_t, t \geq 0)$ deserves the name expectation semigroup, it is a straightforward exercise with the help of the Markov branching property to show that

$$(2.4) \qquad \psi_{t+s}[g](r,v) = \psi_t[\psi_s[g]](r,v), \qquad s, t \geq 0.$$

2.2. Mild form of the neutron transport equation. The reason the choice of notation (2.3) is preemptive is because we can now show how it relates to the aforementioned mild form of the NTE. In order to do so, let us momentarily introduce some more notation: the deterministic evolution $\mathtt{U}_t[g](r,v) = g(r+vt,v)\mathbf{1}_{\{t<\kappa^D_{r,v}\}}, t \geq 0$, where $\kappa^D_{r,v} := \inf\{t > 0 : r + vt \notin D\}$, represents the advection semigroup associated with a single neutron traveling at velocity v from r, killed on exit from D.

Lemma 2.1 ([13]). *Under* (H1) *and* (H2), *for* $g \in L^+_\infty(D \times V)$, *there exist constants* $C_1, C_2 > 0$ *such that* $\psi_t[g]$, *as given in* (2.3), *is uniformly bounded by* $C_1 \exp(C_2 t)$ *for all* $t \geq 0$. *Moreover,* $(\psi_t[g], t \geq 0)$ *is the unique solution, which is bounded in time, to the so-called mild equation (also called a Duhamel solution in the PDE literature):*

$$(2.5) \qquad \psi_t[g] = \mathtt{U}_t[g] + \int_0^t \mathtt{U}_s[(\mathtt{S}+\mathtt{F})\psi_{t-s}[g]]\mathrm{d}s, \qquad t \geq 0,$$

for which (1.2) *holds.*

In (2.5), $\mathtt{S}$ is the corresponding scattering operator with action

$$\mathtt{S}f(r,v) = \sigma_{\mathtt{s}}(r,v)\int_V \Big(f(r,v') - f(r,v)\Big)\pi_{\mathtt{s}}(r,v,v')\mathrm{d}v', \qquad f \in L^+_\infty(D \times V),$$

and $\mathtt{F}$ the fission operator with action

$$\mathtt{F}f(r,v) = \sigma_{\mathtt{f}}(r,v)\int_V f(r,v')\pi_{\mathtt{f}}(r,v,v')\mathrm{d}v' - \sigma_{\mathtt{f}}(r,v)f(r,v), \qquad f \in L^+_\infty(D \times V);$$

see [13] for further details. The connection of the expectation semigroup (2.3) with NTE (1.1) was explored in the recent articles [13, 29] (see also older work in [14, 32]).

2.3. The many-to-one representation. It turns out that there is a second stochastic representation of the unique solution to (2.5). In order to describe it, we need to introduce the notion of a *neutron random walk* (NRW).

A NRW on D is defined by its scatter rates, $\alpha(r,v)$, $r \in D, v \in V$, and scatter probability densities $\pi(r,v,v')$, $r \in D, v, v' \in V$ such that $\int_V \pi(r,v,v')\mathrm{d}v' = 1$ for all $r \in D, v \in V$. Simply, when issued with a velocity v, the NRW will propagate linearly with that velocity until either it exits the domain D, in which case it is killed, or at the random time $T_{\mathtt{s}}$ a scattering occurs, where $\Pr(T_{\mathtt{s}} > t) = \exp\{-\int_0^t \alpha(r+vt,v)\mathrm{d}s\}$ for $t \geq 0$. When the scattering event occurs in position-velocity configuration (r,v), a new velocity v' is selected with probability $\pi(r,v,v')\mathrm{d}v'$. If we denote by $(R,\Upsilon) = ((R_t,\Upsilon_t), t \geq 0)$ the position-velocity of the resulting continuous-time random walk on $D \times V$ with an additional cemetery state $\{\dagger\}$ for when it leaves the domain D, then it is easy to show that (R,Υ) is a Markov process. (Note that neither R nor Υ alone is Markovian.) We call the process (R,Υ) an $\alpha\pi$-NRW. It is worth remarking that when $\alpha\pi$ is given as a single rate function, the density π, and hence the rate α, is uniquely identified by normalizing the given product form $\alpha\pi$ to make it a probability distribution.

To describe the second stochastic representation of (2.5), we shall take

$$(2.6) \qquad \alpha(r,v)\pi(r,v,v') = \sigma_{\mathtt{s}}(r,v)\pi_{\mathtt{s}}(r,v,v') + \sigma_{\mathtt{f}}(r,v)\pi_{\mathtt{f}}(r,v,v'), \qquad r \in D, v, v' \in V.$$

We also need to define

$$(2.7) \qquad \beta(r,v) = \sigma_{\mathtt{f}}(r,v)\left(\int_V \pi_{\mathtt{f}}(r,v,v')\mathrm{d}v' - 1\right) \geq -\sup_{r \in D, v \in V} \sigma_{\mathtt{f}}(r,v) > -\infty,$$

where the uniform lower bound is due to assumption (H1). The following result was established in Lemma 6.1 of [13].

Lemma 2.2 (Many-to-one formula, [13]). *Under the assumptions of Lemma 2.1, we have the second representation*

$$(2.8) \qquad \psi_t[g](r,v) = \mathbf{E}_{(r,v)}\left[\mathrm{e}^{\int_0^t \beta(R_s, \Upsilon_s)\mathrm{d}s} g(R_t, \Upsilon_t)\mathbf{1}_{(t < \tau^D)}\right], \qquad t \geq 0, r \in D, v \in V,$$

where $\tau^D = \inf\{t > 0 : R_t \notin D\}$ and $\mathbf{E}_{(r,v)}$ is the expectation with respect to $\mathbf{P}_{(r,v)}$, which is the law of the $\alpha\pi$-NRW starting from a single neutron with configuration (r,v).

Define $\overline{\beta} := \sup_{r \in D, v \in V} \beta(r,v)$ and note that it is uniformly bounded thanks to (H1) and (H4). Let us now introduce $\mathtt{P}^\dagger := (\mathtt{P}_t^\dagger, t \geq 0)$ for the expectation semigroup of the $\alpha\pi$-neutron random walk with potential β, such as is represented by the semigroup (2.8), but now killed at rate $(\overline{\beta} - \beta)$ or when leaving the domain D. More precisely, for $g \in L_\infty^+(D \times V)$,

$$\mathtt{P}_t^\dagger[g](r,v) = \psi_t[g](r,v)\mathrm{e}^{-\overline{\beta}t}$$
$$= \mathbf{E}_{(r,v)}\left[\mathrm{e}^{\int_0^t (\beta(R_s, \Upsilon_s) - \overline{\beta})\mathrm{d}s} g(R_t, \Upsilon_t)\mathbf{1}_{(t < \tau^D)}\right]$$
$$= \mathbf{E}_{(r,v)}\left[g(R_t, \Upsilon_t)\mathbf{1}_{(t < \mathtt{k})}\right]$$
$$(2.9) \qquad =: \mathbf{E}_{(r,v)}^\dagger[g(R_t, \Upsilon_t)], \qquad t \geq 0, r \in D, v \in V,$$

where

$$\mathtt{k} = \inf\left\{t > 0 : \int_0^t (\overline{\beta} - \beta(R_s, \Upsilon_s))\mathrm{d}s > \mathbf{e}\right\} \wedge \tau^D,$$

and $\mathbf{e}$ is an independent exponentially distributed random variable with mean 1. We will henceforth write $\mathbf{P}^\dagger := (\mathbf{P}_{(r,v)}^\dagger, r \in \bar{D}, v \in V)$ for the associated (sub)probability measures associated to $\mathbf{E}_{(r,v)}^\dagger$, $r \in \bar{D}, v \in V$. The family $\mathbf{P}^\dagger$ now defines a Markov family of probability measures on the path space of the neutron random walk with cemetery state $\{\dagger\}$; note that paths are sent to the cemetery state either when hitting the boundary ∂D or the clock associated to the killing rate $\overline{\beta} - \beta$ rings. Note also that this process is a killed piecewise deterministic Markov process, as introduced in [16].

Remark 2.2. It is worthy of discussion that the proof of Theorem 3.1 in [29] uses the representation (2.9). Moreover, examination of the proof there reveals that, despite the setting of (2.9) forcing a relationship between the value β and the cross-sections in the NTE, this is not needed in the proof. The only requirement is that β be uniformly bounded from above. We will use this fact later on in this paper.

3. Spectral asymptotics. Theorem 5.3 in [13] also gives the leading order behavior of the mild solution of the NTE, mirroring similar results for the setting where the NTE is cited as an abstract Cauchy problem (cf. [29] and [41]); for convenience we reproduce it here.

Theorem 3.1. Suppose that (H1) and (H2) hold. Then, for the semigroup $(\psi_t, t \geq 0)$ identified by (2.5), there exist a $\lambda_* \in \mathbb{R}$, a positive[4] right eigenfunction $\varphi \in L_\infty^+(D \times V)$, and a left eigenmeasure[5] 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 $g \in L_\infty^+(D \times V)$,

$$(3.1) \qquad \langle \tilde{\varphi}, \psi_t[g] \rangle = \mathrm{e}^{\lambda_* t} \langle \tilde{\varphi}, g \rangle \quad (\text{resp.,} \ \psi_t[\varphi] = \mathrm{e}^{\lambda_* t} \varphi), \quad t \geq 0.$$

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

$$(3.2) \qquad \sup_{g \in L_\infty^+(D \times V): \|g\|_\infty \leq 1} \left\| \mathrm{e}^{-\lambda_* t} \varphi^{-1} \psi_t[g] - \langle g, \tilde{\varphi} \rangle \right\|_\infty = O(\mathrm{e}^{-\varepsilon t}) \text{ as } t \to \infty.$$

In parallel with the asymptotic spectral decomposition of the semigroup given in (3.2), [29, 26] also considered analogous stochastic behavior, which we briefly discuss.

An important object that emerges in this context is the process

$$W_t := \mathrm{e}^{-\lambda_* t} \frac{\langle \varphi, X_t \rangle}{\langle \varphi, \mu \rangle}, \qquad t \geq 0,$$

which is a martingale with unit mean under $\mathbb{P}_\mu$. As a nonnegative martingale, it automatically has an almost sure limit, and, from [29], we have the following understanding of its asymptotic behavior.

Theorem 3.2 ([29]). The martingale limit W_∞ is identically zero almost surely if and only if $\lambda_* \leq 0$;, otherwise $(W_t, t \geq 0)$ is $L_2(\mathbb{P}_\mu)$ -convergent with $\mathbb{E}_\mu[W_\infty] = 1$ and $\mathbb{E}_\mu[W_\infty^2] < \infty$.

In essence, we can think of the martingale W as giving us information about the stochastic growth of $\langle g, X_t \rangle$ for the special case that $g = \varphi$. The full picture was given in [26].

Theorem 3.3 ([26]). For all $g \in L_\infty^+(D \times V)$ such that, up to a multiplicative constant, $g \leq \varphi$, under the assumptions of Theorem 3.1, by normalizing $\varphi, \tilde{\varphi}$ such that $\langle \varphi, \tilde{\varphi} \rangle = 1$, we have
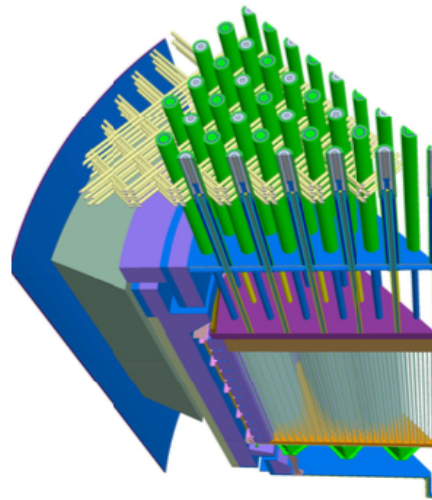
$$(3.3) \qquad \lim_{t \to \infty} \mathrm{e}^{-\lambda_* t} \frac{\langle g, X_t \rangle}{\langle \varphi, \mu \rangle} = \langle g, \tilde{\varphi} \rangle W_\infty,$$

$\mathbb{P}_\mu$-almost surely, for $\mu \in \mathcal{M}(D \times V)$.

---

[4]To be precise, by a positive eigenfunction, we mean a mapping from $D \times V \to (0, \infty)$. This does not prevent it from being valued zero on ∂D, as D is an open bounded, convex domain.

[5]In the forthcoming text we slightly abuse our notation and treat $\tilde{\varphi}$ as both a density and a measure so that $\langle g, \tilde{\varphi} \rangle$ technically means $\int_{D \times V} g(r, v) \tilde{\varphi}(r, v) \mathrm{d}r \mathrm{d}v$.

Figure 1. *The geometry of a nuclear reactor core representing a physical domain D, on which the different cross-sectional values of $\sigma_s, \sigma_f, \pi_s, \pi_f$ are mapped as numerical values.*

The left-eigenfunction $\tilde{\varphi}$ is called the *importance map* and offers a quasi-stationary distribution of radioactive activity (unless $\lambda_* = 0$, in which case it is a stationary profile). We shall refer to the right-eigenfunction as the *heat map* and give a principal representation of neutron density which grows at rate λ_*. In modern nuclear reactor design and safety regulation, it is usually the case that virtual reactor models such as the one seen in Figure 1 are designed such that $\lambda_* = 0$ (in fact, slightly above zero[6]) and the behavior of $\tilde{\varphi}$ and φ within spatial domains on the human scale remain within regulated levels. Existing physics and engineering literature with focus on applications in the nuclear regulation industry has largely been concerned with different numerical methods for estimating the value of the eigenvalue λ_* as well as the eigenfunction φ and eigenmeasure $\tilde{\varphi}(r, \upsilon)\mathrm{d}r\mathrm{d}\upsilon$. Further discussion on this aspect of the NTE can be found in [29, 35, 17, 43, 24, 39, 41, 42, 2, 3, 14, 15].

In this article we look at various existing and new Monte Carlo estimates of the principal quantity λ_*, which opens the door to the much more complicated problem of numerically constructing the eigenfunctions φ and $\tilde{\varphi}$. In particular, we are interested in how Monte Carlo methods based on the recently introduced so-called '*many-to-one*' representation of the solution to (2.5) compare to those based on the classical representation of the underlying physical process. We will exploit the probabilistic perspective that has only recently been developed in the accompanying papers [13, 29, 26]. In particular, we will use the probabilistic perspective to build up complexity analysis of each of the Monte Carlo methods that we consider; considering the balance of computational cost against accuracy.

The remainder of the paper will proceed as follows. In section 4 we show how classical methods use the NBP representation to simulate the lead eigenvalue λ_*, as well as the new

---

[6]As can be seen from [26], reproduced in Theorem 3.3 below, at criticality, the underlying process of fission will die out with probability one, despite the semigroup $(\psi_t, t \geq 0)$ having an asymptotic limit.

NRW representation of $(\psi_t, t \geq 0)$. In sections 5 and 6, we provide a complexity analysis of the two respective methods, which we believe does not currently exist in the literature. From here, a mixture of conclusions can be drawn concerning the relative benefits of the two methods. In section 7 we show how the NRW Monte Carlo method can be adapted using a method of Doob h-transforms, which is a stochastic process version of the method of importance sampling for random variables. This leads to the so-called h-NRW. In section 8 we consider the complexity of the natural Monte Carlo method associated with the h-NRW. Additionally we introduce a candidate class of functions to play the role of h, called *Urts* functions. Finally, in section 9 we discuss the relative benefits of each method and look at some numerical experiments. Almost all of the proofs are deferred to the appendices.

4. Basic Monte Carlo. In Appendix A we outline two algorithmic approaches (Algorithms A.1 and A.2) for the Monte Carlo simulation of an $\alpha\pi$-NRW and a $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}})$-NBP, which were denoted $((\mathbf{r}_t, \mathbf{v}_t), t \leq \mathtt{t}_{\mathrm{end}})$ and $(\mathcal{X}_t, t \geq 0)$, respectively. With either of these simulators in hand, it is possible to build simple statistical estimators of the eigentriple $(\lambda_*, \varphi, \tilde{\varphi})$ based on certain easy limiting procedures that can be observed from Theorem 3.1. For practical reasons noted below, we will in fact focus on estimation of e^{λ_*}, rather than λ_* directly.

Below we introduce such statistical estimators. Notationally we use $\Psi_k[g](t, r, \upsilon)$ to mean one of two statistics that can be built from either an NBP or NRW simulator, where $t \geq 0$, $r \in D$, and $\upsilon \in V$. In the case of NBP, we understand that either

$$(4.1) \qquad \Psi_k[g](t, r, \upsilon) = \Psi_k^{\mathtt{br}}[g](t, r, \upsilon) = \frac{1}{k} \sum_{i=1}^{k} \langle g, \mathcal{X}_t^i \rangle,$$

where $\mathcal{X}^i$ are independent simulations of the NBP described in Algorithm A.2, or

$$(4.2) \qquad \Psi_k[g](t, r, \upsilon) = \Psi_k^{\mathtt{rw}}[g](t, r, \upsilon) = \frac{1}{k} \sum_{i=1}^{k} \mathrm{e}^{\int_0^t \beta(\mathbf{r}_s^i, \mathbf{v}_s^i)\mathrm{d}s} g(\mathbf{r}_t^i, \mathbf{v}_t^i) \mathbf{1}_{(t < \mathtt{t}_{\mathrm{end}}^i)},$$

where $((\mathbf{r}_t^i, \mathbf{v}_t^i), t \leq \mathtt{t}_{\mathrm{end}}^i)$ are independent simulations of the $\alpha\pi$-NRW described in Algorithm A.1. Note that the integral $\int_0^t \beta(\mathbf{r}_s^i, \mathbf{v}_s^i)\mathrm{d}s$ must be computed along simulated paths. In general applications, this might be carried out using standard numerical integration routines, although in nuclear applications, the function β will typically be piecewise constant, and so it can be feasible to use an exact computation methods by breaking the path into lengths along which β is constant.

4.1. Estimating the lead eigenvalue. Taking account of Theorem 3.1, Algorithm A.1 can be used to make the following estimates of λ_*, φ, and $\tilde{\varphi}$, we have in both the NBP and NRW setting that[7]

$$(4.3) \qquad \lim_{t \to \infty} \lim_{k \to \infty} \left(\Psi_k[g](t, r, \upsilon) \right)^{\frac{1}{t}} = \mathrm{e}^{\lambda_*},$$

---

[7]In the case where λ_* is estimated directly, the corresponding estimate would be

$$\lim_{t \to \infty} \lim_{k \to \infty} \frac{1}{t} \log \Psi_k[g](t, r, \upsilon) = \lambda_*.$$

However, this estimate has the undesirable numerical property that for large k, t, there is a positive probability that $\Psi_k[g](t, r, \upsilon) = 0$, and this will cause problems when taking logarithms.

where the limit holds for any $(r, v) \in D \times V$ and bounded measurable $g : D \times V \to [0, \infty)$ such that $\langle g, \tilde{\varphi} \rangle > 0$. Moreover, in the branching setting, if it is a priori known that $\lambda_* > 0$, then thanks to Theorem 3.3, if we take limits in t first, there is no need to involve simulation cycles in the Monte Carlo approach since (conditional on survival of the branching process, which is an event of positive probability)

$$(4.4) \qquad \lim_{t \to \infty} (\Psi_k[g](t, r, v))^{\frac{1}{t}} = \mathrm{e}^{\lambda_*},$$

almost surely. Either way, we can base a Monte Carlo approach on the fact that

$$(\Psi_k[g](t, r, v))^{\frac{1}{t}} \approx \mathrm{e}^{\lambda_*}$$

for k, t sufficiently large and an arbitrary initial configuration $r \in D$, $v \in V$.

4.2. Estimating the lead eigenfunctions. Again, thanks to Theorem 3.1 and Algorithm A.1, for (r, v), $(r_0, v_0) \in D \times V$, in both the NBP and NRW setting we have

$$\lim_{t \to \infty} \lim_{k \to \infty} \frac{\Psi_k[g](t, r, v)}{\Psi_k[g](t, r_0, v_0)} = \frac{\varphi(r, v)}{\varphi(r_0, v_0)},$$

thereby giving an estimate for φ at (r, v) up to a multiplicative constant (consistent for all (r, v) by keeping (r_0, v_0) fixed).

Finally, for fixed $(r, v) \in D \times V$ and bounded measurable $g : D \times V \to [0, \infty)$, if we assume without loss of generality that $\langle \tilde{\varphi}, 1 \rangle = 1$, then

$$(4.5) \qquad \lim_{t \to \infty} \lim_{k \to \infty} \frac{\Psi_k[g](t, r, v)}{\Psi_k[1](t, r, v)} = \langle g, \tilde{\varphi} \rangle$$

for t and k sufficiently large, which gives an estimate for $\tilde{\varphi}$ at (r, v) up to a multiplicative constant. In the setting of the NBP we can again take advantage of Theorem 3.3 and note that, when $\lambda_* > 0$ and extinction does not occur,
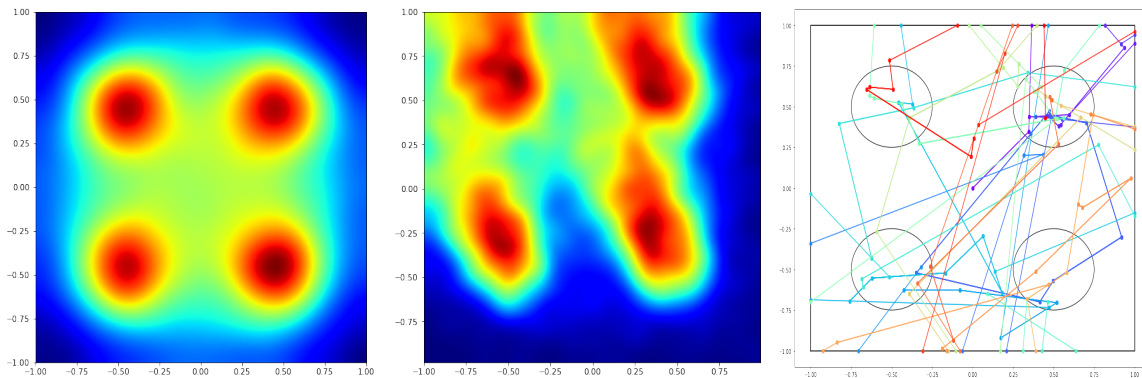
$$(4.6) \qquad \lim_{t \to \infty} \frac{\Psi_1^{\mathtt{br}}[g](t, r, v)}{\Psi_1^{\mathtt{br}}[1](t, r, v)} = \langle g, \tilde{\varphi} \rangle.$$

4.3. Occupation measure eigenfunction estimates at criticality. In the setting that it is a priori known that $\lambda_* = 0$, another approach to estimating φ, and more importantly $\tilde{\varphi}$, involves the use of an occupation measure. Note that on the one hand, for $g \in L_\infty^+(D \times V)$, thanks to Theorem 3.1, it is easy to show that

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \psi_s[g](r, v) \mathrm{d}s = \langle g, \tilde{\varphi} \rangle \varphi(r, v).$$

This implies that one can build an alternative estimate of $\tilde{\varphi}$ by monitoring the empirical occupation of simulations. Specifically, fixing $r_0 \in D$, $v_0 \in V$, and $g \in L_\infty^+(D \times V)$,

$$(4.7) \qquad \lim_{t \to \infty} \lim_{k \to \infty} \lim_{M \to \infty} \frac{1}{tM} \sum_{m=1}^M \Psi_k[g](mt/M, r_0, v_0) = \langle g, \tilde{\varphi} \rangle \varphi(r_0, v_0).$$

Figure 2. *Estimates of $\tilde{\varphi}(r, V) := \int_V \tilde{\varphi}(r, v') dv'$ represented as a "heat map" (left), $\tilde{\varphi}(r, V') := \int_{V'} \tilde{\varphi}(r, v') dv'$ for $V' \subseteq V$ represented as a "heat map" (middle), and some of the contributing underlying branching neutron particles (right). The domain D is rectangular and the cross-sections are constant except on four circular domains, where fissile rates are greater.*

The reader will note that if we take, for example, $g(r, v) = g_{D_0, V_0}(r, v) = \mathbf{1}_{(r \in D_0, v \in V_0)}$, where D_0 is a small region of D and V_0 a small region of V, the quantity

$$\frac{1}{tM} \sum_{m=1}^{M} \Psi_k[g_{D_0, V_0}](mt/M, r, v)$$

(for sufficiently large M, k, t) provides nothing more than a normalized histogram of occupation of the path of the NBP or NRW.

Strictly speaking, a fresh set of Monte Carlo cycles needs to be performed for each "pixel pair" (D_0, V_0). Taking a histogram of occupation across all pixel pairs means that the Monte Carlo simulations on each pixel pair site are highly correlated. For example, in the setting of the NRW, high occupancy in one region of $D \times V$ will necessarily imply lower occupancy in other regions of $D \times V$. Hence, the natural estimator induced by (4.7) has some immediate deficiencies. On the other hand, for the NBP, although the occupancies of individual pixel pairs are highly correlated, the phenomenon of branching ensures that the underling stochastic process explores the space $D \times V$ relatively well over multiple Monte Carlo cycles.

Figure 2 gives an example of the occupation histogram given by $g(r, v) = \mathbf{1}_{(r \in D_0, v \in V)}$, where D_0 varies over small subsets of D. In this example, D is a square, with 4 circular "fuel rods," where the fissile rates are substantially higher than in the surrounding domain. We also give a second example with $g(r, v) = \mathbf{1}_{(r \in D_0, v \in V')}$, where D_0 varies over small subsets of D but $V' \subset V$ is a subset of velocities of V (hence the apparent smear of "heat" in a particular direction in the figure).

5. Complexity analysis of basic NBP Monte Carlo. We turn our attention now to analysing the complexity of some of the algorithms presented in the previous section in the setting that the chosen stochastic process is the physical NBP. Our principal focus will be on the first one, which estimates λ_*. Our interest lies in the rate of convergence of the Monte Carlo estimates from the previous section and the computational cost associated therewith.

We start by looking at the mean-squared convergence of the Monte Carlo estimator suggested by (4.3). Its proof is given in Appendix C.

Theorem 5.1 (NBP Monte Carlo convergence for λ_*). *There exist constants $\kappa_{[i]} := \kappa_{[i]}(g, r, \upsilon)$ such that the following estimates hold for all $t > 0$ and $k \geq 1$:*

(i) *If $\lambda_* = 0$, then*

$$\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\left(\Psi_k^{\mathtt{br}}[g](t,r,\upsilon)\right)^{1/t} - \mathrm{e}^{\lambda_*}\right)^2\right] \leq \frac{\kappa_{[1]} t}{k} + \frac{\kappa_{[0]}}{t^2}\,.$$

(ii) *If $\lambda_* > 0$, then*

$$\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\left(\Psi_k^{\mathtt{br}}[g](t,r,\upsilon)\right)^{1/t} - \mathrm{e}^{\lambda_*}\right)^2\right] \leq \frac{\kappa_{[2]}}{k} + \frac{\kappa_{[0]}}{t^2}\,.$$

(iii) *If $\lambda_* < 0$, then*

$$\mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\left(\Psi_k^{\mathtt{br}}[g](t,r,\upsilon)\right)^{1/t} - \mathrm{e}^{\lambda_*}\right)^2\right] \leq \frac{\kappa_{[3]} \mathrm{e}^{-\lambda_* t}}{k} + \frac{\kappa_{[0]}}{t^2}\,.$$

Let $C_{[0]} := \langle \tilde{\varphi}, g \rangle \varphi(r, \upsilon)$, and $C_{[i]}$ are given in Lemma C.1. If $C_{[0]} \neq 1$, the inequalities hold for all t, k sufficiently large when the constants are chosen, for any $\eta > 0$, to be

$$\kappa_{[0]} := 2\left[\log\left(C_{[0]}\right) \mathrm{e}^{\lambda_*}\right]^2 + \eta,$$
$$\kappa_{[i]} := 2C_{[i]} C_{[0]}^{-2} + \eta, \qquad\qquad i = 1, 2, 3.$$

In particular, this implies that for each (r, υ) and $g \in L_\infty^+(D \times V)$, there exists a sequence $k(t)$ increasing in t such that $k(t) \to \infty$ and such that

(5.1) $$\lim_{t \to \infty} \mathbb{E}_{\delta_{(r,\upsilon)}}\left[\left(\left(\Psi_{k(t)}^{\mathtt{br}}[g](t,r,\upsilon)\right)^{1/t} - \mathrm{e}^{\lambda_*}\right)^2\right] = 0,$$

and the Monte Carlo estimator suggested by (4.3) is asymptotically unbiased with mean-squared convergence.

The above estimates suggest that in the critical and subcritical cases there is a trade-off between increasing t to reduce the second term on the right-hand side of the inequality and an increasing effect on the first term. By also increasing k, it is possible to make the error sufficiently small.

We also want to understand the required computational cost for the associated degree of accuracy in the preceding theorem. To compute the cost of the simulation, we need an indexation for all the particles which have been present in the system.

Suppose that $\mathcal{U}$ is a labeling system that uniquely identifies (or "names") neutrons that ever come into existence in the NBP (for example, one could use, e.g., Ulam Harris labeling; cf. [33]). For each $u \in \mathcal{U}$, write $(r_t^{(u)}, \upsilon_t^{(u)})_{b_u \leq t < d_u}$ for the path of the particle u, where $0 \leq b_u < d_u \leq \infty$ stand for the birth and death times of u. Recall that $r^{(u)}$ is piecewise linear;

write $b_u = t_0^{(u)} < t_1^{(u)} < \cdots$ for the jump times of $v_t^{(u)}$ (the moments when scatterings occur). For $f, g \in L_\infty^+(D \times V)$, define the cost per simulation by

$$C_t[f, g] := \sum_{u \in \mathcal{U}} g\left(r_{b_u}^{(u)}, v_{b_u}^{(u)}\right) \mathbf{1}_{(b_u \leq t)} + \sum_{u \in \mathcal{U}} \sum_{i \geq 1} f\left(r_{t_i^{(u)}}^{(u)}, v_{t_i^{(u)}}^{(u)}\right) \mathbf{1}_{(b_u \leq t, t_i^{(u)} \leq t)}$$

for $t \geq 0$.

The major part of the MC simulation suggested by (4.3) consists of simulating the variables relevant to the scattering and fission events. The random functional $(C_t[f, g], t \geq 0)$, with an appropriate choice of f, g, can then be used to evaluate the cost (if we assume the computational effort of simulating any random variable is counted as one unit) of a single simulation with a time horizon t. For instance, if we take $f = \mathbf{1}$ and $g = \mathbf{0}$, then $C_t[\mathbf{1}, \mathbf{0}]$ counts the total number of scattering events up to time t, which can be used as a rough estimate of the CPU time of the program; on the other hand, taking $f = \mathbf{0}, g = \mathbf{1}$ yields an estimate for the memory cost as $C_t[\mathbf{0}, \mathbf{1}]$ counts the number of particles that have appeared before t.

The process $(C_t[f, g], t \geq 0)$ is a pure jump increasing process whose Doob–Meyer decomposition can be written as $C[f, g] = A + M$, where M is a martingale with respect to the natural filtration of X and A is the bounded variation compensator of $C[f, g]$. The latter can be easily computed by noting that rates at which scattering and fission occur are given by the cross-sections, and it is straightforward to derive

$$(5.2) \qquad A_t = \int_0^t \langle \sigma_\mathbf{s} \pi_\mathbf{s}[f] + \sigma_\mathbf{f} \pi_\mathbf{f}[g], X_s \rangle \mathrm{d}s, \qquad t \geq 0,$$

where for $(r, v) \in D \times V$ and $h \in L_\infty^+(D \times V)$,

$$(5.3)\ \pi_\mathbf{s}[h](r, v) = \int_V h(r, v') \pi_\mathbf{s}(r, v, v') \mathrm{d}v' \quad \text{and} \quad \pi_\mathbf{f}[h](r, v) = \int_V h(r, v') \pi_\mathbf{f}(r, v, v') \mathrm{d}v'.$$

Such computations are not unfamiliar in the setting of branching Markov processes, superprocesses, or fragmentation processes. See, for example, [22, 36, 5]. We are led very easily to the following estimate of simulation cost per sample as a simple application of Theorem 3.1.

Doob's maximal inequality for the representation $C[f, g] = M + A$ also gives, for example,

$$(5.4) \qquad \mathbb{E}_{\delta_{(r,v)}}\left[\sup_{s \leq t} (C_s[f, g] - A_s)^2\right] \leq 4 \int_0^t \psi_s[\sigma_\mathbf{s} \pi_\mathbf{s}[f^2] + \sigma_\mathbf{f} \pi_\mathbf{f}[g^2]](r, v) \mathrm{d}s, \qquad t \geq 0.$$

This shows that the mean-squared error of the growth of the cost function $C[f, g]$ relative to the compensator A mimics the growth of

$$\mathbb{E}_{\delta_{(r,v)}}[C_t[f, g]] = \mathbb{E}_{\delta_{(r,v)}}[A_t] = \int_0^t \psi_s[\sigma_\mathbf{s} \pi_\mathbf{s}[f] + \sigma_\mathbf{f} \pi_\mathbf{f}[g]](r, v) \mathrm{d}s, \qquad t \geq 0.$$

Lemma 5.2 (NBP expected simulation cost per sample). *For $f, g \in L_\infty^+(D \times V)$ we have the following:*

(i) *If* $\lambda_* = 0$, *then, as* $t \to \infty$,

(5.5) $$\mathbb{E}_{\delta_{(r,v)}}[C_t[f,g]] \sim \langle \sigma_{\mathtt{s}}\pi_{\mathtt{s}}[f] + \sigma_{\mathtt{f}}\pi_{\mathtt{f}}[g], \tilde{\varphi}\rangle \varphi(r,v) t =: \kappa_{[4]} t \,.$$

(ii) *If* $\lambda_* > 0$, *then, as* $t \to \infty$,

(5.6) $$\mathbb{E}_{\delta_{(r,v)}}[C_t[f,g]] \sim \langle \sigma_{\mathtt{s}}\pi_{\mathtt{s}}[f] + \sigma_{\mathtt{f}}\pi_{\mathtt{f}}[g], \tilde{\varphi}\rangle \varphi(r,v) \frac{\mathrm{e}^{\lambda_* t}}{\lambda_*} = \kappa_{[4]} \frac{\mathrm{e}^{\lambda_* t}}{\lambda_*} \,.$$

(iii) *If* $\lambda_* < 0$, *then, as* $t \to \infty$, $\mathbb{E}_{\delta_{(r,v)}}[C_t[f,g]] \nearrow \mathbb{E}_{\delta_{(r,v)}}[C_\infty[f,g]] =: \kappa_{[5]} < \infty$.

Of course, if we run k cycles of Algorithm A.2, the cost of these cycles is then a k-multiple of that of a single cycle as stipulated in Lemma 5.2. It is also worth noting that both $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}[f] + \sigma_{\mathtt{f}}\pi_{\mathtt{f}}[g] < \infty$ and $\sigma_{\mathtt{s}}\pi_{\mathtt{s}}[f^2] + \sigma_{\mathtt{f}}\pi_{\mathtt{f}}[g^2] < \infty$ for $f, g \in L_\infty^+(D \times V)$.

We conclude with the following consequence of Theorem 5.1 and Lemma 5.2. Specifically, we want to identify the optimal choice of k, t to minimize total computational cost for a given level of accuracy. To this end, we define

$$\mathtt{Cost}(k,t) = k \times \sup_{f,g \in L_\infty^+(D \times V) : \|f\|_\infty, \|g\|_\infty \leq 1} C_t[f,g].$$

Theorem 5.3 (NBP Monte Carlo complexity). *There is a choice of k, t such that*

(5.7) $$\mathbb{E}_{\delta_{(r,v)}}\left[\left(\left(\Psi_k^{\mathtt{br}}[g](t,r,v)\right)^{1/t} - \mathrm{e}^{\lambda_*}\right)^2\right] \leq \varepsilon^2,$$

with the following total simulation cost:

(i) *If* $\lambda_* = 0$, *then* (5.7) *holds in the limit as* $\varepsilon \to 0$ *with*

$$\mathbb{E}_{\delta_{(r,v)}}[\mathtt{Cost}(k,t)] \leq C\varepsilon^{-4}$$

for some C, and the optimal choice of k, t is given by $t = \sqrt{2\kappa_{[0]}}\varepsilon^{-1}$ and $k = \kappa_{[0]}\kappa_{[1]}^{-1}t^3$.

(ii) *If* $\lambda_* > 0$, *then* (5.7) *holds in the limit as* $\varepsilon \to 0$ *with*

$$\mathbb{E}_{\delta_{(r,v)}}[\mathtt{Cost}(k,t)] \leq C \exp\left(\lambda_* \sqrt{\kappa_{[0]}}\varepsilon^{-1}\right)$$

for some C, and the optimal choice of k, t corresponds to $t \approx \sqrt{\kappa_{[0]}}\varepsilon^{-1}$ and $k \approx \lambda_* \kappa_{[0]}^{1/2}\kappa_{[2]}\varepsilon^{-3}/2$.

(iii) *If* $\lambda_* < 0$, *then* (5.7) *holds in the limit as* $\varepsilon \to 0$ *with*

$$\mathbb{E}_{\delta_{(r,v)}}[\mathtt{Cost}(k,t)] \leq C \exp\left(|\lambda_*| \sqrt{\kappa_{[0]}}\varepsilon^{-1}\right)$$

for some C, and the optimal choice of k, t corresponds to $t \approx \sqrt{\kappa_{[0]}}\varepsilon^{-1}$ and $k \approx |\lambda_*|\kappa_{[0]}^{1/2}\kappa_{[3]}\varepsilon^{-3}\exp(|\lambda_*|t)/2$.

Proof. (i) Following Theorem 5.1 and Lemma 5.2, it follows that we want to solve the optimization problem:

$$\text{minimize} \quad tk \quad \text{subject to} \quad \frac{t\kappa_{[1]}}{k} + \frac{\kappa_{[0]}}{t^2} \leq \varepsilon^2, \quad k, t > 0.$$

Introducing a Lagrangian variable $\xi \geq 0$ to penalize the constraint, we see that we expect the minimum to occur when

$$k + \frac{\kappa_{[1]}\xi}{k} - 2\kappa_{[0]}\xi t^{-3} = 0, \qquad t - \frac{t\kappa_{[1]}\xi}{k^2} = 0.$$

Rearranging the second expression, and noting that $t > 0$, we deduce $k^2 = \kappa_{[1]}\xi$, and hence, substituting back into the first expression, that $k = \kappa_{[1]}\kappa_{[0]}^{-1}t^3$. Substituting this into the constraint, we deduce that $t = \sqrt{2\kappa_{[0]}}\varepsilon^{-1}$, and the expression for the optimal cost follows immediately.

(ii) Similarly to the previous case, we need to solve the problem:

$$\text{minimize} \quad e^{\lambda_* t}k \quad \text{subject to} \quad \frac{\kappa_{[2]}}{k} + \frac{\kappa_{[0]}}{t^2} \leq \varepsilon^2, \quad k, t > 0.$$

Proceeding as above to introduce a Lagrangian penalization, and optimizing over t and k, we deduce that $2\xi\kappa_{[0]} = \lambda_* kt^3 e^{\lambda_* t} = 2\kappa_{[0]}\kappa_{[2]}^{-1}k^2 e^{\lambda_* t}$, and hence $k = \kappa_{[2]}\kappa_{[0]}^{-1}\lambda_* t^3/2$. Substituting into the constraint, we deduce that the optimal t satisfies $t^3\varepsilon^2 - \kappa_{[0]}t - 2\kappa_{[0]}\lambda_*^{-1} = 0$. It follows for ε sufficiently small that $t \leq \sqrt{\kappa_{[0]}}\varepsilon^{-1} + \eta$ for some $\eta > 0$. Recalling that $\kappa_{[0]}$ is already defined only up to some small constant, we can assume $\eta = 0$ and absorb an ε^{-3} term into the same constant. The remainder of the argument then follows directly as above.

(iii) The case where $\lambda_* < 0$ follows with an almost identical argument; in this case the optimization problem is

$$\text{minimize} \quad k \quad \text{subject to} \quad \frac{\kappa_{[3]}e^{-\lambda_* t}}{k} + \frac{\kappa_{[0]}}{t^2} \leq \varepsilon^2, \quad k, t > 0,$$

and the answer follows from an essentially identical set of steps to (ii). ∎

6. Complexity Analysis of Basic NRW Monte Carlo. Next we turn our attention to the setting of the Monte Carlo simulation based on the NRW estimator—in other words, via the representation (2.8). The picture in this setting is typically less competitive than in the the NBP estimator setting; however, this approach will offer additional flexibility in terms of possible algorithmic approaches to finding λ_*. Recall that $\overline{\beta} := \sup_{r \in D, v \in V} \beta(r, v)$ and introduce also $\underline{\beta} := \inf_{r \in D, v \in V} \beta(r, v)$.

Theorem 6.1 (NRW Monte Carlo convergence for λ_*). *For all* $\lambda_* \in \mathbb{R}$ *and* $g \in L_\infty^+(D \times V)$ *there exist* $\tilde{\kappa}_{[1]} > 0$ *and* $\lambda_1 \in \mathbb{R}$ *such that*

$$\mathbf{E}_{(r,v)}\left[\left((\Psi_k^{\mathtt{rw}}[g](t,r,v))^{1/t} - e^{\lambda_*}\right)^2\right] \leq \tilde{\kappa}_{[1]}\frac{1}{k}e^{(\lambda_1 - 2\lambda_*)t} + \frac{\kappa_{[0]}}{t^2}$$

as $t \to \infty$. *Moreover,* $(2\lambda_* \vee (\lambda_* + \underline{\beta})) \leq \lambda_1 \leq \lambda_* + \overline{\beta}$.

It is immediately obvious from the above result that, since there is no setting in which the NRW Monte Carlo offers better asymptotic accuracy than the NBP Monte Carlo algorithm does. The reason for this is quite simple. It is a straightforward consequence of the previously observed fact that a large majority of simulations for the NRW Monte Carlo estimator will return zero contribution to the estimate on account of the indicator in (2.8).

The constant $\tilde{\kappa}_{[1]}$ can be computed using convergence results in the appendix (see (D.6)) as $2\mathrm{e}^{2\lambda_*}\left(\varphi_1(r,v)\langle\tilde{\varphi}_1,g^2\rangle - (\varphi(r,v)\langle\tilde{\varphi},g\rangle)^2\mathbf{1}_{\lambda_1=2\lambda_*}\right)/C_{[0]}^2$. Here $\varphi_1,\tilde{\varphi}_1$ are analogues of $\varphi,\tilde{\varphi}$ for a related semigroup specified in the appendices.

In terms of computational cost, it is again a simple exercise to identify the analogue of Lemma 5.2. In this setting we need only define the cost of each simulation by the function

$$(6.1) \qquad C_t[f] = \sum_j f(R_{t_j}, \Upsilon_{t_j})\mathbf{1}_{(t_j \leq t)}, \qquad t \geq 0,$$

where $0 \leq t_0 < t_1 < \cdots$ are the scatter times of the process (R, Υ) and $f \in L_\infty^+(D \times V)$. As in the setting of Lemma 5.2, we can write $C[f] = A + M$, where M is a martingale with respect to the natural filtration of (R, Υ), and A is the bounded variation compensator of $C[f]$. As before, the scattering rates of (R, Υ) together with standard computations allow us to write, for $t \geq 0$,

$$(6.2) \qquad A_t = \int_0^{t \wedge \tau^D} \alpha\pi[f](R_s, \Upsilon_s)\mathrm{d}s, \qquad t \geq 0,$$

where $\pi[h]$ is defined similarly to (5.3).

On account of the fact that

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

we can use the assumptions (H1) and (H2) to deduce that

$$\sup_{r \in D, v \in V} \alpha(r,v) \leq \overline{c}$$

for constant $0 < \overline{c} < \infty$. We can think of the trajectory of process (R, Υ) as a series of randomly placed sticks laid end to end in the domain D. Suppose that $S_1, S_2, \ldots, S_N$ are the successive lengths of these sticks, where N is the last stick which touches the boundary of D (with the understanding that $N = \infty$ if this does not happen). The estimate above ensures that we can construct a sequence of independent and identically distributed random variables, which are exponentially distributed with rate 1, say $\mathbf{e}_1, \mathbf{e}_2, \ldots$, such that $S_i \geq \upsilon_{\mathtt{min}}\mathbf{e}_i/\overline{c}$, for $i = 1, \ldots, N$. It now follows that the probability that the length of any given stick is bounded below by $\mathrm{diam}(D) = \sup\{|x - y| : x, y \in \partial D\}$ is uniformly bounded below by $\exp(-\mathrm{diam}(D)\overline{c}/\upsilon_{\mathtt{min}})$. Accordingly, it follows that N is upper bounded by Γ, where Γ is a geometrically distributed random variable with parameter $\exp(-\mathrm{diam}(D)\overline{c}/\upsilon_{\mathtt{min}})$. Moreover, we also have that

$$\tau^D \leq \Gamma\mathrm{diam}(D)/\upsilon_{\mathtt{min}},$$

so that, in particular, $\sup_{r \in D, v \in V} \mathbb{E}_{\delta_{(r,v)}}[\tau^D] < \infty$. We thus have the following result for the computational cost.

Lemma 6.2 (NRW expected simulation cost). *For $f \in L_\infty^+(D \times V)$*

$$\lim_{t \to \infty} \mathbf{E}_{(r,v)}[C_t[f]] = \mathbf{E}_{(r,v)} \left[\int_0^{\tau^D} \alpha \pi[f](R_s, \Upsilon_s) \mathrm{d}s \right] < \infty.$$

We should of course note again that if there are k cycles in the Monte Carlo simulation, then the cost grows no faster than proportional to k, irrespective of the time horizon.

As with the previous section, we conclude with a corollary, which gives the complexity of the Monte Carlo algorithm. On this occasion

$$\mathtt{Cost}(k,t) = k \times \sup_{f \in L_\infty^+(D \times V): \|f\|_\infty \leq 1} C_t[f]$$

for $k \geq 1$, $t \geq 0$.

Corollary 6.3 (NRW Monte Carlo complexity). *There is a choice of k, t such that*

(6.3)
$$\mathbb{E}_{\delta_{(r,v)}} \left[\left((\Psi_k^{\mathtt{rw}}[g](t, r, v))^{1/t} - \mathrm{e}^{\lambda_*} \right)^2 \right] \leq \varepsilon^2.$$

Moreover, (6.3) holds in the limit as $\varepsilon \to 0$ with

$$\mathbb{E}_{\delta_{(r,v)}}[\mathtt{Cost}(k,t)] \leq C \exp\left((\lambda_1 - 2\lambda_*)\sqrt{\kappa_{[0]}} \varepsilon^{-1} \right)$$

for some $C > 0$, and the optimal choice of k, t corresponds to $t \approx \sqrt{\kappa_{[0]}} \varepsilon^{-1}$ and $k \approx (\lambda_1 - 2\lambda_*)\sqrt{\kappa_{[0]}} \tilde{\kappa}_{[1]} \varepsilon^{-3} \exp((\lambda_1 - 2\lambda_*)t)/2$.

Proof. From Lemma 6.2 we have, for all $t > 0$, $\mathbb{E}_{\delta_{(r,v)}}[\mathtt{Cost}(k,t)] \leq C'k$ for some constant C', and hence the objective is to

$$\text{minimize} \quad k \quad \text{subject to} \quad \frac{\tilde{\kappa}_{[1]} \mathrm{e}^{(\lambda_1 - 2\lambda_*)t}}{k} + \frac{\kappa_{[0]}}{t^2} \leq \varepsilon^2, \quad k, t > 0.$$

This is essentially identical to the minimization carried out in (iii) of Theorem 5.3, and the result follows. ∎

7. Importance sampling with NRW Monte Carlo. One common method for improving the efficiency of Monte Carlo estimation schemes is through the use of importance sampling. Recalling, for example, (4.2), the term $\mathrm{e}^{\int_0^t \beta(\mathbf{r}_s^i, \mathbf{v}_s^i)\mathrm{d}s} \mathbf{1}_{\{t < \mathbf{t}_{\mathrm{end}}^i\}}$ in the estimator can be interpreted as the weight of the ith particle. In particular, considering the simple case where $g \equiv 1$, one wants to apply importance sampling to reduce the variance of this particle weight as much as possible by simulating paths from a suitably chosen probability distribution.

In this context, we will show that the correct way to think about importance sampling in the case of the NTE is through the use of Doob's h-transform. In this section, we show how

to h-transform the NRW, and we prove that under mild assumptions this gives an unbiased estimate of the semigroup $(\psi_t; t \geq 0)$.

Appealing to the Markov property together with Theorem 3.1, suppose we write $\mathcal{G}_t = \sigma((R_s, \Upsilon_s), s \leq t)$ for $t \geq 0$. Then, for $0 \leq s \leq t < \infty$,

$$\mathbf{E}_{(r,v)}\left[e^{-\lambda_* t} e^{\int_0^t \beta(R_u, \Upsilon_u)\mathrm{d}u} \frac{\varphi(R_t, \Upsilon_t)}{\varphi(r,v)} \mathbf{1}_{(t<\tau^D)} \middle| \mathcal{G}_s \right]$$

$$= \mathbf{1}_{(s<\tau^D)} e^{-\lambda_* s} e^{\int_0^s \beta(R_u, \Upsilon_u)\mathrm{d}u}$$

$$\times \mathbf{E}_{(r',v')}\left[e^{-\lambda_*(t-s)} e^{\int_0^{t-s} \beta(R_u, \Upsilon_u)\mathrm{d}u} \frac{\varphi(R_{t-s}, \Upsilon_{t-s})}{\varphi(r,v)} \mathbf{1}_{(t-s<\tau^D)} \right]_{(r',v')=(R_s, \Upsilon_s)}$$

$$= \mathbf{1}_{(s<\tau^D)} e^{-\lambda_* s} e^{\int_0^s \beta(R_u, \Upsilon_u)\mathrm{d}u} e^{-\lambda_*(t-s)} \psi_{t-s}[\varphi](R_s, \Upsilon_s)/\varphi(r,v)$$

$$(7.1) \quad = \mathbf{1}_{(s<\tau^D)} e^{-\lambda_* s} e^{\int_0^s \beta(R_u, \Upsilon_u)\mathrm{d}u} \varphi(R_s, \Upsilon_s)/\varphi(r,v).$$

Together with the boundedness of φ and the Markov property, this tells us that

$$e^{-\lambda_* t} e^{\int_0^t \beta(R_u, \Upsilon_u)\mathrm{d}u} \frac{\varphi(R_t, \Upsilon_t)}{\varphi(r,v)} \mathbf{1}_{(t<\tau^D)}, \qquad t \geq 0,$$

is a martingale.

As such we can use this martingale to effect a change of measure on the path space of $(R, \Upsilon) = ((R_t, \Upsilon_t), t \geq 0)$ via the relation

$$(7.2)$$

$$\mathbf{P}^\varphi_{(r,v)}(A, t < \tau^D) = \mathbf{E}_{(r,v)}\left[e^{-\lambda_* t} e^{\int_0^t \beta(R_s, \Upsilon_s)\mathrm{d}s} \frac{\varphi(R_t, \Upsilon_t)}{\varphi(r,v)} \mathbf{1}_{(A, t<\tau^D)} \right] = e^{-\lambda_* t} \frac{\psi_t[\varphi \mathbf{1}_A](r,v)}{\varphi(r,v)}$$

for all $(r,v) \in D \times V$ and $A \in \sigma((R_s, \Upsilon_s), s \leq t)$. The theory of Doob h-transforms dictates that the process (R, Υ) under $\mathbf{P}^\varphi_{(r,v)}$, $r \in D$, $v \in V$, is again a Markov process; see [12]. It is conservative only if we can prove that $\mathbf{P}^\varphi_{(r,v)}(\tau^D = \infty) = 1$ for all $(r,v) \in D \times V$, which is to say, the process (R, Υ) is trapped in D. However this is easily verified from (7.2) by taking A to be the entire sample space and invoking the martingale property to deduce that $\mathbf{P}^\varphi_{(r,v)}(t < \tau^D) = 1$ for all $t \geq 0$. Moreover, since $\tilde{\varphi}, \varphi \in L_\infty^+(D \times V)$, we can normalize $\tilde{\varphi}$ and φ so that $\langle \tilde{\varphi}, \varphi \rangle = 1$; for bounded and measurable $g : \bar{D} \times V \to [0, \infty)$ we have

$$\lim_{t \to \infty} \mathbf{E}^\varphi_{(r,v)}[g(R_t, \Upsilon_t)] = \lim_{t \to \infty} e^{-\lambda_* t} \frac{\psi_t[\varphi g](r,v)}{\varphi(r,v)} = \langle g, \varphi\tilde{\varphi} \rangle,$$

indicating that $\varphi\tilde{\varphi}$ is the stationary distribution of (R, Υ) under $\mathbf{P}^\varphi_{(r,v)}$, $r \in D$, $v \in V$.

As is usual with such space-time changes of measure, the action of its infinitesimal generator, say L_φ, can be heuristically calculated by applying the h-transformation to the generator of the process (R, Υ), say L, with potential $\beta - \lambda_*$ under $\mathbf{P}_{(r,v)}$, $(r,v) \in \bar{D} \times V$. Given the discussion in section 2.3, we have the action of the generator L satisfying

$$(7.3) \qquad \mathsf{L}f(r,v) = v \cdot \nabla f(r,v) + \alpha(r,v) \int_V (f(r,v') - f(r,v)) \pi(r,v,v')\mathrm{d}v'$$

boilerplate on the left margin:

for all f in the domain of L, written Dom(L); e.g., we can take f as continuous in r and differentiable in v. Hence, on $D \times V$,

$$
\begin{aligned}
\mathtt{L}_\varphi f(r,v) &= (\mathtt{L} + \beta - \lambda_*)^\varphi f(r,v) \\
&= \frac{1}{\varphi(r,v)} \mathtt{L}(\varphi f)(r,v) + \beta f(r,v) - \lambda_* f(r,v) \\
&= v \cdot \nabla f(r,v) + \frac{v \cdot \nabla \varphi(r,v)}{\varphi(r,v)} f(r,v) + \beta f(r,v) - \lambda_* f(r,v) \\
&\quad + \frac{\alpha(r,v)}{\varphi(r,v)} \int_V (f(r,v') - f(r,v))\, \varphi(r,v')\pi(r,v,v')\mathrm{d}v' \\
&\quad + f(r,v)\frac{\alpha(r,v)}{\varphi(r,v)} \int_V (\varphi(r,v') - \varphi(r,v))\, \pi(r,v,v')\mathrm{d}v' \\
&= v \cdot \nabla f(r,v) + \frac{(\mathtt{L} + \beta - \lambda_*)\varphi(r,v)}{\varphi(r,v)} f(r,v) \\
&\quad + \alpha(r,v) \int_V (f(r,v') - f(r,v))\, \frac{\varphi(r,v')}{\varphi(r,v)}\pi(r,v,v')\mathrm{d}v'
\end{aligned}
\tag{7.4}
$$

for all bounded and continuous $f \in \mathrm{Dom}(\mathtt{L})$. From Theorem 3.1, we can interpret (3.1) as heuristically meaning that $(\mathtt{L} + \beta - \lambda_*)\varphi = 0$. Taking this on face value, we get

$$
\mathtt{L}_\varphi f(r,v) = v \cdot \nabla f(r,v) + \int_V (f(r,v') - f(r,v))\, \alpha(r,v)\frac{\varphi(r,v')}{\varphi(r,v)}\pi(r,v,v')\mathrm{d}v'.
$$

In other words, $\mathtt{L}_\varphi$ corresponds to an $(\alpha_\varphi, \pi_\varphi)$-NRW, where, for $r \in D$ and $v \in V$,

$$
\alpha_\varphi(r,v) = \alpha(r,v) \int_V \frac{\varphi(r,v')}{\varphi(r,v)}\pi(r,v,v')\mathrm{d}v' \quad \text{and} \quad \pi_\varphi(r,v,v') = \frac{\varphi(r,v')\pi(r,v,v')}{\int_V \varphi(r,v'')\pi(r,v,v'')\mathrm{d}v''}.
$$

An interesting observation in light of this change of measure is that, appealing to (7.2), by choosing $g \in L_\infty^+(D \times V)$, we have that

$$
\psi_t[g](r,v) = \mathbf{E}_{(r,v)}^\varphi \left[\mathrm{e}^{\lambda_* t} \frac{\varphi(R_t, \Upsilon_t)^{-1}}{\varphi(r,v)^{-1}} g(R_t, \Upsilon_t) \right].
\tag{7.5}
$$

This leads to a thought experiment. If we knew φ and λ_* and we take $g \equiv \varphi$, then the term inside the expectation on the right-hand side is simply a constant. As a consequence, $\psi_t[\varphi](r,v)$ can be computed exactly with a single Monte Carlo trajectory. In particular, we would be interested in a guess that would inform a Doob h-transformed process that would produce estimates for $\psi_t[g]$, which involves averaging a low-variance term that is close to the quantity inside the expectation in (7.5). The problem with this thought experiment is, of course, that we are aiming to generate Monte Carlo estimates of $\psi_t[g](r,v)$ with a view to estimating the quantities λ_* and φ.

The idea behind the thought experiment is not lost, however. Suppose instead we were to make an educated guess for φ and λ_*. In particular, we would be interested in a guess that would inform a Doob h-transformed process that we could simulate with relatively few (if any) paths that leave the domain D. The next theorem formalizes this idea (which also gives a rigorous basis to the discussion above by setting $h = \varphi$).

Theorem 7.1. *Suppose that $h \in L_\infty^+(D \times V)$ such that $h \geq 0$ on $D \times V$, $h|_{\partial D^+} = 0$, and*

$$(7.6) \qquad \inf_{r \in D} \int_V h(r, v')\mathrm{d}v' > 0.$$

Define

$$(7.7) \qquad \mathrm{J}h(r, v) = \alpha(r, v) \int_V \left(h(r, v') - h(r, v)\right) \pi(r, v, v')\mathrm{d}v'$$

for $D \times V$. Suppose that we denote by $(T_i, i \geq 0)$ the scattering times of (R, Υ) with $T_0 = 0$. Moreover, let $N_t = \sup\{i : T_i \leq t\}$. Then

$$(7.8) \qquad \left.\frac{\mathrm{d}\mathbf{P}_{(r,v)}^h}{\mathrm{d}\mathbf{P}_{(r,v)}}\right|_{\sigma((R_s, \Upsilon_s), s \leq t)} := \exp\left(-\int_0^t \frac{\mathrm{J}h(R_s, \Upsilon_s)}{h(R_s, \Upsilon_s)}\mathrm{d}s\right) \prod_{i=1}^{N_t} \frac{h(R_{T_i}, \Upsilon_{T_i})}{h(R_{T_i}, \Upsilon_{T_{i-1}})} \mathbf{1}_{(t < \tau^D)}$$

characterizes an NRW whose expectation semigroup is given by

$$(7.9) \qquad \psi_t^h[g](r, v) = \mathrm{U}_t[g](r, v) + \int_0^t \mathrm{U}_s[\mathrm{J}_h \psi_{t-s}^h[g]](r, v)\mathrm{d}v',$$

where the scattering operator is given by

$$(7.10) \qquad \mathrm{J}_h g(r, v) = \int_V \left(g(r, v') - g(r, v)\right) \alpha(r, v) \frac{h(r, v')}{h(r, v)} \pi(r, v, v')\mathrm{d}v'$$

on D. Moreover,

$$\psi_t[g](r, v)$$
$$(7.11) \qquad = \mathbf{E}_{(r,v)}^h\left[\exp\left(\int_0^t \frac{\mathrm{J}h(R_s, \Upsilon_s)}{h(R_s, \Upsilon_s)} + \beta(R_s, \Upsilon_s)\mathrm{d}s\right) \prod_{i=1}^{N_t} \frac{h(R_{T_i}, \Upsilon_{T_{i-1}})}{h(R_{T_i}, \Upsilon_{T_i})} g(R_t, \Upsilon_t)\mathbf{1}_{(t < \tau^D)}\right]$$

for $t \geq 0$, and if we additionally assume that

$$(7.12) \qquad \inf_{r \in D, v \in V} \lim_{s \to \kappa_{r,v}^D} \frac{|v|(\kappa_{r,v}^D - s)}{h(r + vs, v)} > 0,$$

where we recall that $\kappa_{r,v}^D := \inf\{t > 0 : r + vt \notin D\}$, and then the term $\mathbf{1}_{(t < \tau^D)}$ in (7.8) and (7.11) can be removed and (R, Υ) under $\mathbf{P}_{(r,v)}^h$, $r \in D, v \in V$, is conservative.

Remark 7.1. We also note from (7.9), i.e., the expectation semigroup of (R, Υ) under $\mathbf{P}_{(r,v)}^h$, $r \in D, v \in V$, that it is formally associated to the generator

$$(7.13) \qquad \mathrm{L}_h g(r, v) = v \cdot \nabla g(r, v) + \int_V \left(g(r, v') - g(r, v)\right) \alpha(r, v) \frac{h(r, v')}{h(r, v)} \pi(r, v, v')\mathrm{d}v'$$

on $D \times V$.

Suppose, moreover, that we write $\mathtt{T} = v \cdot \nabla$. If we further assume that $\mathtt{T}h$ is well defined, for example, one could typically assume that $\mathtt{T}$ is defined weakly in $L_2(D \times V)$, then we can also write the term

$$(7.14) \quad \frac{h(r + vt, v)}{h(r, v)} = \exp\left(\log h(r + vt, v) - \log h(r, v)\right) = \exp\left(\int_0^t \frac{\mathtt{T}h(r + vs, v)}{h(r + vs, v)} \mathrm{d}s\right).$$

As such, a piecewise application of (7.14) allows us to convert the right-hand side of (7.11) into

$$(7.15) \quad \psi_t[g](r, v) = h(r, v)\mathbf{E}_{(r,v)}^h\left[\exp\left(\int_0^t \frac{\mathtt{L}h(R_s, \Upsilon_s)}{h(R_s, \Upsilon_s)} + \beta(R_s, \Upsilon_s)\mathrm{d}s\right) \frac{g(R_t, \Upsilon_t)}{h(R_t, \Upsilon_t)} \mathbf{1}_{(t < \tau^D)}\right].$$

Similarly the change of measure (7.8) can be identified alternatively as

$$(7.16) \quad \left.\frac{\mathrm{d}\mathbf{P}_{(r,v)}^h}{\mathrm{d}\mathbf{P}_{(r,v)}}\right|_{\sigma((R_s, \Upsilon_s), s \leq t)} := \exp\left(-\int_0^t \frac{\mathtt{L}h(R_s, \Upsilon_s)}{h(R_s, \Upsilon_s)} \mathrm{d}s\right) \frac{h(R_t, \Upsilon_t)}{h(r, v)} \mathbf{1}_{(t < \tau^D)}.$$

It is worth remarking that the representation of the semigroup $(\psi_t, t \geq 0)$ in (7.15) takes on a more convenient form than in (7.11). However, technically speaking, it requires more assumptions on the function h.

Proof of Theorem 7.1. Let us introduce $\psi_t^h[g](r, v) = \mathbf{E}_{(r,v)}^h[g(R_t, \Upsilon_t)\mathbf{1}_{(t < \tau^D)}]$ for $t \geq 0$, $r \in D$, $v \in V$, and $g \in L_\infty(D \times V)$, and note that we may also write

$$\psi_t^h[g](r, v) = \mathbf{E}_{(r,v)}\left[\exp\left(-\int_0^t \frac{\mathtt{J}h(R_s, \Upsilon_s)}{h(R_s, \Upsilon_s)} \mathrm{d}s\right) \prod_{i=1}^{N_t} \frac{h(R_{T_i}, \Upsilon_{T_i})}{h(R_{T_i}, \Upsilon_{T_{i-1}})} g(R_t, \Upsilon_t)\mathbf{1}_{(t < \tau^D)}\right].$$

Using standard arguments centered around the Markov property applied at the first scattering time, it is easy to show that $(\psi_t^h, t \geq 0)$ is an expectation semigroup.

Next, by conditioning on the first step of the NRW under $\mathbf{P}_{(r,v)}$, we get

$$\psi_t^h[g](r, v)$$
$$= \exp\left(-\int_0^t \alpha(r + vs, v) + \frac{\mathtt{J}h(r + vs, v)}{h(r + vs, v)} \mathrm{d}s\right) g(r + vt, v)\mathbf{1}_{(t < \kappa_{r,v}^D)}$$
$$+ \int_0^t \mathbf{1}_{(s < \kappa_{r,v}^D)} \alpha(r + vs, v) \exp\left(-\int_0^s \alpha(r + vu, v) + \frac{\mathtt{J}h(r + vu, v)}{h(r + vu, v)} \mathrm{d}u\right)$$
$$(7.17)$$
$$\int_V \psi_{t-s}^h[g](r + vs, v') \frac{h(r + vs, v')}{h(r + vs, v)} \pi(r + vs, v, v') \mathrm{d}v' \mathrm{d}s.$$

Next note that from the assumptions (H1) and (H2) and (7.6),

$$\frac{\mathtt{J}h(r, v)}{h(r, v)} + \alpha(r, v) = \frac{\alpha(r, v)}{h(r, v)} \int_V h(r, v')\pi(r, v, v') \geq \frac{1}{h(r, v)} \underline{\alpha\pi} \int_V h(r, v')\mathrm{d}v',$$

where $\underline{\alpha} = \inf_{r \in D, v \in V} \alpha(r, v) > 0$ and $\underline{\pi} = \inf_{r \in D, v, v' \in V} \pi(r, v, v') > 0$. Together with assumption (7.12) and (7.6), this ensures that

$$\lim_{t \uparrow \kappa_{r,v}^D} \int_0^t \left\{ \frac{Jh(r + vs, v)}{h(r + vs, v)} + \alpha(r + vs, v) \right\} \mathrm{d}s = \infty$$

for all $r \in D$, $v \in D$. Since α is uniformly bounded from above, it follows that the indicators in (7.17) are not necessary.

Now appealing to Lemma 1.2, Chapter 4 in [21], we can remove exponential potential terms at the expense of introducing an additive potential term. After application of the aforementioned lemma, we have

$$\psi_t^h[g](r, v) = g(r + vt, v)$$
$$+ \int_0^{t \wedge \kappa_{r,v}^D} \alpha(r + vs, v) \int_V \psi_{t-s}^h[g](r + vs, v') \frac{h(r + vs, v')}{h(r, v)} \pi(r + vs, v, v') \mathrm{d}v'$$

(7.18)
$$- \int_0^{t \wedge \kappa_{r,v}^D} \left(\alpha(r + vs, v) + \frac{Jh(r + vs, v)}{h(r + vs, v)} \right) \psi_{t-s}^h[g](r + vs, v) \mathrm{d}s.$$

Noting again that

(7.19)
$$\alpha(r, v) + \frac{Jh(r, v)}{h(r, v)} = \alpha(r, v) \int_V \frac{h(r, v')}{h(r, v)} \pi(r, v, v') \mathrm{d}v'$$

and recalling the definition (7.10), plugging (7.19) back into (7.18), we get

$$\psi_t^h[g](r, v) = g(r + vt, v) + \int_0^{t \wedge \kappa_{r,v}^D} J_h \psi_{t-s}^h[g](r + vs, v) \mathrm{d}v'$$

(7.20)
$$= \mathtt{U}_t[g](r, v) + \int_0^t \mathtt{U}_s[J_h \psi_{t-s}^h[g]](r, v) \mathrm{d}v'.$$

Equation (7.20) together with the semigroup property shows that the change of measure (7.8) makes (R, Υ) under $\mathbf{P}_{(r,v)}^h$, $r \in D, v \in V$, the law of an NRW whose expectation semigroup is formally associated to the infinitesimal generator (7.13).

To verify conservativeness, it suffices to take $g = 1$ in (7.20) and note that, since $J1 = 0$, then $\psi_t^h[g](r, v) = 1$ is a solution to (7.20) for all t. Similarly to the mild equation (2.5), the solution to (7.20) is unique (thanks to a simple application of Grönwall's lemma); see, for example, [29]. Thus we have conservativeness. ∎

As a small side remark, we note that the representation in Theorem 7.1 is an alternative form of Feynman–Kac representation which works with multiplicative potentials instead of exponential additive potentials. Examples where this has been used for other Markov systems are discussed in [10, 31].

8. Complexity analysis of h-NRW Monte Carlo and the shape of h. Under the assumptions of Theorem 7.1 (not necessarily including (7.12)) and Remark 7.1, in particular through

the representation (7.15), we see a third approach to simulating λ_*. More precisely, with the estimate (4.3) we have the alternative estimator

$$(8.1) \qquad \Psi_k[g](t,r,\upsilon) = \Psi_k^{h\text{-}\mathtt{rw}}[g](t,r,\upsilon) = h(r,\upsilon)\frac{1}{k}\sum_{i=1}^{k}\exp\left(\int_0^t \frac{\mathtt{L}h(\mathbf{r}_s^i,\mathbf{v}_s^i)}{h(\mathbf{r}_s^i,\mathbf{v}_s^i)} + \beta(\mathbf{r}_s^i,\mathbf{v}_s^i)\mathrm{d}s\right)$$
$$\frac{g(\mathbf{r}_t^i,\mathbf{v}_t^i)}{h(\mathbf{r}_t^i,\mathbf{v}_t^i)}\mathbf{1}_{(t<\mathtt{t}_{\mathrm{end}}^i)},$$

where $((\mathbf{r}_t^i,\mathbf{v}_t^i), t\geq 0)$ are independent copies of an $\alpha^h\pi^h$-NRW, with

$$(8.2) \qquad \alpha^h(r,\upsilon)\pi^h(r,\upsilon,\upsilon') = \alpha(r,\upsilon)\frac{h(r,\upsilon')}{h(r,\upsilon)}\pi(r,\upsilon,\upsilon'), \qquad r\in D, \upsilon,\upsilon'\in V.$$

(See Algorithm A.3 in the appendix.)

Following the analysis in section 6, we use similar calculations to control the variance and cost of the estimator associated to (8.1). For variance control, we have the following result, whose conclusion does not depend on the sign of λ_*.

Theorem 8.1 (h-NRW Monte Carlo convergence for λ_*). *Suppose that the assumptions of Theorem 7.1, including (7.12), and of Remark 7.1 hold. Additionally, suppose*

$$(8.3) \qquad \sup_{r\in D, \upsilon\in V} \frac{\mathtt{L}h(r,\upsilon)}{h(r,\upsilon)} < \infty.$$

Then for $g\in L_\infty^+(D\times V)$ such that $g\leq h$, there exists a constant $\tilde{\kappa}_{[2]} > 0$ such that

$$(8.4) \qquad \mathbf{E}_{(r,\upsilon)}^h\left[\left(\left(\Psi_k^{h\text{-}\mathtt{rw}}[g](t,r,\upsilon)\right)^{1/t} - \mathrm{e}^{\lambda_*}\right)^2\right] \leq \frac{\tilde{\kappa}_{[2]}\mathrm{e}^{(\lambda_2-2\lambda_*)t}}{k} + \frac{\kappa_{[0]}}{t^2},$$

as $t\to\infty$, where $\lambda_* + \underline{\varsigma} \leq \lambda_2 \leq \lambda_* + \overline{\varsigma}$ and

$$(8.5) \qquad -\infty \leq \inf_{r\in D, \upsilon\in V} \frac{(\mathtt{L}+\beta)h(r,\upsilon)}{h(r,\upsilon)} =: \underline{\varsigma} \leq \overline{\varsigma} := \sup_{r\in D, \upsilon\in V} \frac{(\mathtt{L}+\beta)h(r,\upsilon)}{h(r,\upsilon)}.$$
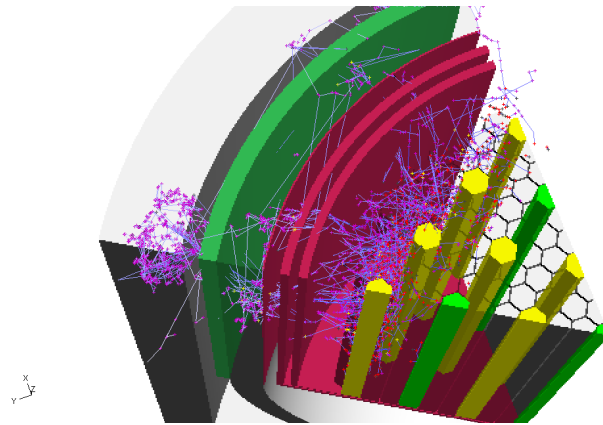
Next we turn our attention to the cost analysis in the spirit of Lemma 6.2.

Lemma 8.2 (h-NRW Expected simulation cost). *Suppose that the conditions of Theorem 8.1 are met and additionally that $\underline{\varsigma} > -\infty$. Then, for $f\in L_\infty^+(D\times V)$ and some constant $K_f > 0$,*

$$\limsup_{t\to\infty} \mathrm{e}^{-(\lambda_*-\underline{\varsigma})t}\mathbf{E}_{(r,\upsilon)}^h[C_t[f]] < K_f,.$$

Moreover, $\mathbf{E}_{(r,\upsilon)}^h[C_t[f]]$ grows at most linearly in t if either $\underline{\varsigma} = \lambda_*$ or if the following are satisfied:

(i) *$f\in L_\infty^+(D\times V)$ satisfies $\pi^h[f](r,\upsilon) \geq c_0 > 0$ for all $(r,\upsilon)\in D\times V$;*
(ii) *there exist $\Omega_0 \subseteq D\times V$, $p_0 > 0$, and $\delta > 0$ such that*

    (a) *$\int_{(r,\upsilon')\in\Omega_0} \pi^h(r,\upsilon,\upsilon')\,\mathrm{d}\upsilon' \geq p_0$;*
    (b) *$\alpha^h(r+\upsilon s,\upsilon) \leq \delta^{-1}$ for all $s\in[0,\delta]$, for all $(r,\upsilon)\in\Omega_0$.*

Figure 3. An example of a slice through a mock nuclear reactor design in which some neutron path simulations are depicted. The reactor core displays symmetry.

Remark 8.1. Heuristically speaking, the upper bound in Theorem 8.1 is optimized by choosing $h = \varphi$. Indeed, in that setting, $(\mathtt{L} + \beta)\varphi = \lambda_* \varphi$ on $L_2(D \times V)$. Hence, assuming this is a pointwise equality, it follows that $\lambda_2 = 2\lambda_*$, matching the lower bound for λ_2, thereby minimizing the upper bound in (8.4). Note also that in the same setting $\varsigma = \lambda_*$, and hence the growth of the expected cost is linear.

Of course, in practice the function φ is generally unknown and is a quantity that we want to learn through Monte Carlo methods. In the context of importance sampling, knowing φ would be equivalent to knowing the optimal weights for the simulated random variable.

The sufficient condition (7.12) is strongly indicative of linear behavior of h for directional derivatives towards the physical boundary of D. These features should be built into any first estimate of h. We also note that the main theorem in underlying theory, i.e., Theorem 3.1, requires that the domain D is convex. In other contexts, where the NTE is interpreted as an abstract Cauchy problem on a Banach space, it sometimes additionally assumed that the cross-sections on $D \times V$ are also piecewise continuous, or even more simply, piecewise constant. This is a very natural assumption for the application of nuclear reactor core design. See, for example, Figure 3, which shows a typical simulation in a slice of a nuclear reactor core which has convex symmetries.

Let us assume momentarily that D is a convex polyhedron. That is to say, it is a domain bounded by a finite number of tangent planes. Below we will construct two possible families of choices of h. Both are built using a combination of hyperplanes.

Example 8.3 (Urts functions). For each $v \in V$, one possible definition we can work with is

$$h(r, v) = c \times \operatorname{dist}_{v/|v|}(r, \partial D),$$

where, for $\omega \in \mathbb{S}_2$, $\operatorname{dist}_\omega(r, \partial D)$ is the distance from r to the boundary of D along the ray $(r + \omega t, t \geq 0)$ and $c > 0$. Note that this is a piecewise linear function in r for fixed v. Indeed, we can write this more precisely as
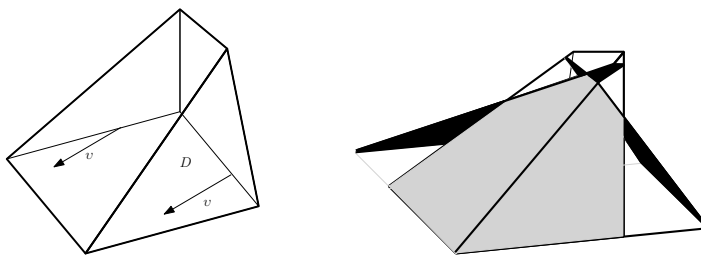
Figure 4. *On the left $h(r,v) = c_1\mathrm{dist}_{v/|v|}(r, \partial D)$, and on the right is the case of an Urts function $h(r,v) = c_1\mathrm{dist}_{v/|v|}(r, \partial D) \wedge c_2\mathrm{dist}_{-v/|v|}(r + r_v, \partial D)$. In each figure, the "base" represents the domain D, and the height represents the value of the function h.*

$$h(r,v) = c \times \inf\left\{ t > 0 : r + \frac{v}{|v|}t \in \partial D \right\} = c|v|\inf\{s > 0 : r + vs \in \partial D\} = c|v|\kappa_{r,v}^D.$$

Note that

$$(8.6) \qquad \frac{|v|(\kappa_{r,v}^D - s)}{h(r + vs, v)} = \frac{\kappa_{r,v}^D - s}{c\kappa_{r+vs,v}^D} = \frac{1}{c}$$

and hence (7.12) is satisfied.

In two dimensions this appears to take the form of the minimum of hyperplanes touching only the sides of ∂D that v is directed towards (i.e., only $r \in \partial D$ for which $v \cdot \mathbf{n}_r > 0$). See the function on the left in Figure 4.

Recalling (7.10), this proposal for h thus imposes an instantaneous repulsive rate for r close to boundary points $r' \in \partial D$ such that $v \cdot \mathbf{n}_{r'} > 0$. However, if r is positioned far from the aforementioned boundary points but close to one of the boundary points $r' \in \partial D$ such that $v \cdot \mathbf{n}_{r'} < 0$, we should also expect a milder repulsion from this boundary. We can thus adjust our proposed approximation to φ by choosing it to be such that it receives a gentler repulsion from the boundary which is diametrically opposed to the direction of travel. We thus propose a first estimate of φ taking the piecewise linear form

$$(8.7) \qquad h(r,v) = \min\{c_1\mathrm{dist}_{v/|v|}(r, \partial D)\,, \; c_2\mathrm{dist}_{-v/|v|}(r + r_v, \partial D)\},$$

where r_v can be dependent on v or just taken as a fixed constant and $c_1, c_2 > 0$ are constants. The role of r_v is to provide the "gentle repulsion" away from the boundaries as alluded to above. The second term in the minimum is a simple measure of how far the boundary might be after scattering, although naturally more sophisticated choices could be made, at the expense of more complexity in subsequent calculations.

We call this family of approximations for φ Urts functions, named after the teepee-like dwellings of the Mongolian Siberian Tsaatan tribes which are similarly shaped when D is a two-dimensional convex polyhedral domain; see Figure 4. As before, it is easy to verify that $\inf_{r\in D} \int_V h(r, v')\mathrm{d}v' > 0$.

Unfortunately, neither of the previous two examples necessarily respect the condition (8.3). To see why, suppose we consider the setting of an Urts function with $c_1 = c_2 = 1$. For (8.3)

to hold, we would need in particular that, for $r \in \partial D$ and $v \in V$ such that $v \cdot \mathbf{n}_r > 0$ (i.e., the velocity direction of v is pointing *out* of the domain D), the limit

$$
\lim_{s \to 0} \frac{\mathtt{T}h(r - sv/|v|, v) + \mathtt{J}h(r - sv/|v|, v)}{h(r - sv/|v|, v)}
$$

$$
= \lim_{s \to 0} \left(\alpha(r, v) \int_V \frac{h(r - sv/|v|, v')}{h(r - sv/|v|, v)} \pi(r, v, v') \mathrm{d}v' - \frac{|v|}{h(r - sv/|v|, v)} \right) - \alpha(r, v)
$$

$$
(8.8) \qquad = \lim_{s \to 0} \left(\alpha(r, v) \int_V \frac{1}{s} h(r - sv/|v|, v') \pi(r, v, v') \mathrm{d}v' - \frac{|v|}{s} \right) - \alpha(r, v)
$$

needs to exist and be finite. Observe that in the last step we used the form of the function h to deduce that $h(r - sv/|v|, v) = s$ for s small, and $r \in \partial D$ and $v \in V$ such that $v \cdot \mathbf{n}_r > 0$. Note that this is potentially problematic because for v' in the integral on the right-hand side above such that $v' \cdot \mathbf{n}_r < 0$, we have $h(r, v') > 0$, and hence the factor $1/s$ means that, in the limit as $s \downarrow 0$, if it exists, can be explosive. Indeed, on the one hand, for s sufficiently close to zero,

$$
\lim_{s \to 0} \int_{\{v' \cdot \mathbf{n}_r > 0\}} \frac{h(r - sv/|v|, v')}{h(r - sv/|v|, v)} \pi(r, v, v') \mathrm{d}v' = \lim_{s \to 0} \int_{\{v' \cdot \mathbf{n}_r > 0\}} \frac{(v \cdot \mathbf{n}_r)|v'|}{(v' \cdot \mathbf{n}_r)|v|} \pi(r, v, v') \mathrm{d}v' < \infty.
$$

On the other hand, for v' such that $v' \cdot \mathbf{n}_r < 0$, we have that $h(r, v') > 0$ and hence, for s sufficiently small,

$$
\int_{\{v' \cdot \mathbf{n}_r < 0\}} \frac{h(r - sv/|v|, v')}{h(r - sv/|v|, v)} \pi(r, v, v') \mathrm{d}v' > \frac{C}{s}
$$

for some constant $C > 0$, which explodes as $s \to 0$. On the other hand, so long as the supremum in (8.8) is uniformly bounded above (i.e., the limit could be $-\infty$), then (8.3) is respected.

Example 8.4 (lifted Urts functions). The previous discussion indicated a possible way of building an Urts function such that the requirement (8.3) is satisfied. Another way to do this is to violate the zero boundary condition that $h(r, v) = 0$ when $v \cdot \mathbf{n}_r > 0$ and to take the Urts function of Example 8.3 and add a small constant to it. That is to say,

$$
h(r, v) = \varepsilon + \min\{c_1 \mathrm{dist}_{v/|v|}(r, \partial D), \, c_2 \mathrm{dist}_{-v/|v|}(r + r_v, \partial D)\}
$$

for some $0 < \varepsilon \ll 1$ and $c_1, c_2 > 0$. For a lifted Urts function, the condition (7.12) is thus not satisfied and, as a consequence, not all of the h-NRW will survive until the chosen time horizon.

Lemma 8.5. *Suppose that the domain $D \subset \mathbb{R}^2$ is bounded by a regular polygon or circle, and h is one of the urts functions described above. Then the conditions of* (ii) *of Lemma 8.2 are satisfied; in particular, the cost $\mathbf{E}^h_{(r,v)}[C_t[f]]$ will grow at most linearly whenever f is uniformly bounded below by a strictly positive constant.*

Proof. For simplicity, we consider the case where $D = [-L, L] \times [-L, L]$. The proof in other cases is similar.

Fix $\eta > 0$, and define

$$\Omega_1 := ([-L, L - \eta] \times [-L, L - \eta]) \times \{v \in V : v/|v| = (\cos(\theta), \sin(\theta)), \theta \in [\pi/8, 3\pi/8]\},$$
$$\Omega_2 := ([-L + \eta, L] \times [-L, L - \eta]) \times \{v \in V : v/|v| = (\cos(\theta), \sin(\theta)), \theta \in [5\pi/8, 7\pi/8]\},$$
$$\Omega_3 := ([-L + \eta, L] \times [-L + \eta, L]) \times \{v \in V : v/|v| = (\cos(\theta), \sin(\theta)), \theta \in [9\pi/8, 11\pi/8]\},$$
$$\Omega_4 := ([-L, L - \eta] \times [-L + \eta, L]) \times \{v \in V : v/|v| = (\cos(\theta), \sin(\theta)), \theta \in [13\pi/8, 15\pi/8]\}$$

and $\Omega_0 := \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4$. Then it follows from each of the definitions of the Urts functions that h is bounded below on Ω_0, and bounded above on $D \times V$. Hence we deduce that (ii).(a) holds. Similarly, the fact that h is bounded below on Ω_0 and above on $D \times V$ guarantees that (ii).(b) holds.

Note in particular that (ii).(a) guarantees also condition (i) of Lemma 8.2 when f is bounded below, uniformly away from zero. ∎

We conclude with the corresponding version of Corollary 6.3.

Corollary 8.6 (h-NRW Monte Carlo complexity). *Suppose all conditions of Theorem 8.1 are satisfied, and we are in the case of Lemma 8.2 where $\mathbf{E}^h_{(r,v)}[C_t[f]]$ grows at most linearly. Then there is a choice of k, t such that*

$$(8.9) \qquad \mathbb{E}_{\delta_{(r,v)}}\left[\left(\left(\Psi_k^{\mathtt{h-rw}}[g](t, r, v)\right)^{1/t} - e^{\lambda_*}\right)^2\right] \leq \varepsilon^2.$$

Moreover, (8.9) holds in the limit as $\varepsilon \to 0$ with

$$\mathbb{E}_{\delta_{(r,v)}}[\mathtt{Cost}(k, t)] \leq C \exp\left((\lambda_2 - 2\lambda_*)\sqrt{\kappa_{[0]}}\varepsilon^{-1}\right)$$

for some $C > 0$, and the optimal choice of k, t corresponds to $t \approx \sqrt{\kappa_{[0]}}\varepsilon^{-1}$ and $k \approx (\lambda_2 - 2\lambda_*)\sqrt{\kappa_{[0]}}\tilde{\kappa}_{[2]}\varepsilon^{-3}\exp((\lambda_2 - 2\lambda_*)t)/2$.

The proof follows in a similar manner to Corollary 6.3.

9. Discussion and numerical comparison of methods.

9.1. Relation to existing Monte Carlo methods.
There are many interesting connections between the methods we have discussed in this paper and existing Monte Carlo methods discussed in the literature. Of particular relevance for our discussion (and in particular in comparison with the NBP and NRW methods described) are various Monte Carlo methods which have been proposed for simulating from the quasi-stationary distribution (QSD) of a given Markov process. We highlight some relevant strands fo research below:

- A common approach to simulate from the QSD of a killed Markov process is to use *Fleming–Viot* particle systems. In this approach, N particles are simulated independently from a (killed) Markov process, up to the first time one of the particles is killed, and the killed particle is replaced by a copy of a randomly chosen surviving particle. See, e.g., [8, 25, 9]. Under fairly weak conditions, this process converges to the QSD as the simulation time, and the number of particles goes to infinity. Such methods are not immediately applicable in the case of the NTE, however, since they generally require a condition on the particles that the killing times for independent particles

cannot coincide. In general this is not true for the NTE. In addition, existing methods do not handle branching, although this can be incorporated into the NRW setting by increasing the killing rate when the branching rate is bounded above. Below we will also discuss how other approaches based on interacting particle systems may also be applied in our setting.

- Another, potentially interesting, alternative approach to simulation from the QSD involves stochastic approximation approaches [4, 38]. Again, these methods look to simulate the QSD of a killed process. In this setting, rather than simulating many particles and using other particles to resurrect killed particles, killed particles are resurrected at some point in their history. In this manner, a single, potentially infinite trajectory is sampled, and a quantity based on the occupation measure (cf. section 4.3) is used to estimate the QSD. Existing results are not directly applicable to our setting, either because the results are in discrete time or are diffusion models in continuous time (see [47]); however, it would appear to be an interesting question for future work to see if these methods can be fruitfully adapted to the NTE setting.

- There are a number of recent papers in the Markov chain Monte Carlo (MCMC) literature which look to simulate from a given measure, typically motivated by applications in Bayesian statistics. In these applications, and unlike the NTE, the measure is generally known (at least up to a normalizing constant), and the challenge is to construct Markov processes which realize the given measure as their (quasi-)stationary distribution. In [44] the measure is constructed as the QSD of a diffusion model and interacting particle methods applied to simulate from the QSD. Other approaches (see [46, 44]) highlight the benefit of using piecewise deterministic Markov processes (PDMPs) as an efficient method to rapidly converge to the QSD, and show how to efficiently constrict PDMPs which have a desired stationary measure. Since the NTE is an example of a PDMP, it would be interesting to understand the extent to which numerical methods for the NTE can borrow from, or inspire ideas for, these novel MCMC methods.

9.2. Towards improved Monte Carlo methods for the NTE.
The results of the previous section (in particular Remark 8.1) suggests that extremely efficient methods for estimating λ_* can be developed if the eigenfunction φ is known. Unfortunately, finding the optimal φ is generally at least as hard as finding the optimal λ_*, so this is not immediately helpful. However, importance sampling can still be highly effective in reducing the variance of the numerical scheme even when the function φ is not known. The key question is how to construct informative functions h which might approximate φ well.

Moreover, in practical implementations, the naive approach suggested in Theorem 5.3 is generally not amenable in practice. Considering the asymptotic results given here, one first observes that it is heavily desirable to be in the $\lambda_* = 0$ regime, which has linear growth, rather than the cases when $\lambda_* \neq 0$, which both have exponential growth in complexity for a given level of accuracy. It follows that it is numerically beneficial to try to get the system to an equilibrium state, for example, by introducing additional killing or branching (although the optimal rate is of course unknown, it could, for example, be estimated numerically from simulations). Moreover, numerical methods which try to mollify excess particle deaths or births through

repopulating from the existing population can also be effective at keeping the system close to equilibrium. Such methods lead naturally to considering interacting particle systems (e.g., [18]) and sequential Monte Carlo (SMC) methods (see [20]). In most industrial-standard software,[8] numerical implementations follow a modified version of classical SMC methods known as power iteration ([37, Chap. 6.III], [6]) where particles are randomly selected from a "birth store," representing the current particle population, with unused particles being stored for possible later use. The eigenvalue is estimated by adjusting the birthrate in a manner that tries to keeps the population constant. We are not aware of any rigorous theoretical justification or analysis of these methods, but it would seem that they would bear significant similarity to standard SMC methods, and we would expect similar theoretical behavior.

The benefits of combining SMC methods with importance sampling has been investigated by, e.g., [49]. It is in this combination that we expect the results in sections 7 and 8 to be of most benefit. Specifically, we would anticipate using SMC methods to sample/resample from a population of particles, which themselves undergo motion according to an h-transformed version of the NRW. We aim to write more about this in forthcoming work.

In practice, in the nuclear engineering literature, a large range of different variance reduction methods are used. One class of variance reduction methods appears to be closely related to the methods introduced in sections 7 and 8. In the nuclear engineering literature, these variance reduction methods are called *zero-variance* calculations; see [11], building on earlier work stretching back to the 1940s. Most of the nuclear engineering literature considers zero-variance schemes in the context of shielding problems, where the goal is to estimate the neutron density at a detection site, based on a fixed source of particles, and where fissile behavior is typically not dominant; [11] is a notable exception to this, although, unlike the current paper, this paper considers the case of generational growth, not time-dependent growth, and the presentation is not mathematically rigorous. To the best of our knowledge, there is no work in the nuclear engineering literature which considers the zero-variance solutions to time-dependent criticality problems, although these are current areas of active research; see, for example, [50, 23, 40].

A further difficulty that arises in the use of h-transformed motion for the estimation of λ_* is that the estimates can be dominated by large-deviation effects. Specifically, although the estimates remain unbiased, for large t a substantial contribution to the expectation in (7.11) comes from rare particles with large weights (the exponential, product, and indicator terms in the expression). For example, [30], in a slightly different setting, connects the limiting value λ to a maximization problem, which in our setting reads as

$$\lambda_* \geq \lim_{T \to \infty} \left\{ \mathbb{E}^h \left[\beta(R_T, \Upsilon_T) \right] - \frac{1}{T} \mathbb{E}^h \left[\log \left(\left. \frac{\mathrm{d}\mathbb{P}^h}{\mathrm{d}\mathbb{P}} \right|_T \right) \right] \right\}.$$

Moreover, the optimal choice of $h = \varphi$ will attain equality in this expression. By manipulation of this expression, it should be possible to develop iterative schemes which improve h to get better estimates of λ_*. In a different setting, this is the approach taken recently by [28]. We leave the complete discussion of how these methods may be implemented in the case of the NTE to future work.

---

[8]For example, MCNP [48], MONK [45], Serpent [34], and Tripoli-4 [7].

In the rest of this section we provide some basic numerical experiments[9] that demonstrate the convergence rates observed in the previous sections. We concentrate on an analytically tractable one-dimensional case, as well as a two-dimensional setting, which is not analytically tractable. We conclude with some discussion of the relative strengths and weaknesses of the methods.

9.3. One-dimensional slab reactor. We consider a particular example where D is an open interval; cf. [11, 50]. Up to some linear transformation of the variables, we can always assume that $D = \{(r_1, r_2, r_3) : -L < r_1 < L\}$ for some fixed $(r_2, r_3) \in \mathbb{R}^2$ and $L \in (0, \infty)$. Clearly in this case $(\psi_t, t \geq 0)$ is merely a function of the first coordinate. We can therefore suppress the dependency on the second and third coordinates and reduce the problem entirely to its one-dimensional form. In other words, we can reduce the entire problem to the setting $D = (-L, L)$.

The interest of this example lies in that computations of the leading eigenvalue and the associated eigenfunctions in this case are tractable so that we can compare numerical results with theoretical values. Throughout this section, we assume that

$$(9.1) \qquad V = \{-v_0, v_0\}, \quad \sigma_{\mathtt{s}}(r, v) \equiv \sigma_{\mathtt{s}}, \quad \sigma_{\mathtt{f}}(r, v) \equiv \sigma_{\mathtt{f}},$$
$$\pi_{\mathtt{s}}(r, v, v') = \delta_{-v}(\mathrm{d}v'), \quad \pi_{\mathtt{f}}(r, v, \mathrm{d}v') = 2\delta_v(\mathrm{d}v'),$$

where $v_0, \sigma_{\mathtt{s}}, \sigma_{\mathtt{f}} \in (0, \infty)$. Note that, as a small technical matter, because the velocity space V consists of just two elements, we need to replace $\pi_{\mathtt{s}}(r, v, v')\mathrm{d}v'$ by $\pi_{\mathtt{s}}(r, v, \mathrm{d}v')$ and a similar replacement for $\pi_{\mathtt{f}}$. Observe that we have $\sigma = \sigma_{\mathtt{s}} + \sigma_{\mathtt{f}}$ and the two measures $\pi_{\mathtt{f}}(r, v, \mathrm{d}v')$ and $\pi_{\mathtt{s}}(r, v, \mathrm{d}v')'$ are singular with respect to one another. Note that under assumption (9.1) the operator $\mathtt{A} = \mathtt{T} + \mathtt{S} + \mathtt{F}$ takes the following simple form:

$$(9.2) \quad \begin{cases} \mathtt{T}f(r, v) = v \cdot \nabla f(r, v) & \text{(transport)}, \\[2mm] \mathtt{S}f(r, v) = \sigma_{\mathtt{s}}(f(r, -v) - f(r, v)) & \text{(scattering)}, \\[2mm] \mathtt{F}f(r, v) = \sigma_{\mathtt{f}}(2f(r, v) - f(r, v)) = \sigma_{\mathtt{f}}f(r, v) & \text{(fission)} \end{cases}$$

for all $f \in L_2^+(D \times V)$ and $(r, v) \in D \times V$.

The main condition of Theorem 3.1 requires that $\sigma_{\mathtt{f}}(2-1) = \sigma_{\mathtt{f}} > 0$, which is automatically the case. Moreover, Theorem 8.1 of [13] allows us to otherwise identify the eigentriple $(\lambda_*, \varphi, \tilde{\varphi})$ via the relations

$$(9.3) \qquad \langle \tilde{\varphi}, \mathtt{A}g \rangle = \lambda_* \langle \tilde{\varphi}, g \rangle \text{ and } \langle f, \mathtt{A}\varphi \rangle = \lambda_* \langle f, \varphi \rangle, \qquad f, g \in L_2^+(D \times V),$$

where the eigenfunctions are identified this way only uniquely within $L_2^+(D \times V)$ (which is sufficient for our purposes to identify the eigenfunctions in $L_\infty^+(D \times V)$).

---

[9]The numerical simulations have all been implemented in Python. The code can be downloaded from people.bath.ac.uk/mapamgc/.

Proposition 9.1. *The leading eigenvalue λ_* and the associated eigenfunctions $\varphi, \tilde{\varphi}$ as specified in Theorem* 3.1 *are given as follows.*

(i) *If $\theta := v_0/2L\sigma_{\mathbf{s}} > 1$, let x_* be the unique nonnegative solution to the equation*

$$(9.4) \qquad \frac{\sinh(x)}{x} = \theta.$$

Then the leading eigenvalue is given as

$$(9.5) \qquad \lambda_* = \sigma_{\mathbf{f}} - \sigma_{\mathbf{s}} - \sqrt{\sigma_{\mathbf{s}}^2 + \left(\tfrac{v_0 x_*}{2L}\right)^2},$$

and we have

$$\varphi(r, v) = \phi(r)\mathbb{1}_{\{v=v_0\}} + \phi(-r)\mathbb{1}_{\{v=-v_0\}}, \quad \tilde{\varphi}(r, v) = \phi(-r)\mathbb{1}_{\{v=v_0\}} + \phi(r)\mathbb{1}_{\{v=-v_0\}},$$

where

$$(9.6) \qquad \phi(r) = \frac{\sinh\left(\tfrac{1}{2}x_*(1 - \tfrac{r}{L})\right)}{\sinh(\tfrac{1}{2}x_*)}, \qquad r \in (-L, L).$$

(ii) *If $\theta := v_0/2L\sigma_{\mathbf{s}} = 1$, then the leading eigenvalue is given as*

$$(9.7) \qquad \lambda_* = \sigma_{\mathbf{f}} - 2\sigma_{\mathbf{s}},$$

and we have

$$\varphi(r, v) = \phi(r)\mathbb{1}_{\{v=v_0\}} + \phi(-r)\mathbb{1}_{\{v=-v_0\}}, \quad \tilde{\varphi}(r, v) = \phi(-r)\mathbb{1}_{\{v=v_0\}} + \phi(r)\mathbb{1}_{\{v=-v_0\}},$$

where

$$(9.8) \qquad \phi(r) = 1 - \frac{r}{L}, \qquad r \in (-L, L).$$

(iii) *If $\theta := v_0/2L\sigma_{\mathbf{s}} < 1$, then let x_* be the smallest nonnegative solution to the equation*

$$(9.9) \qquad \frac{\sin(x)}{x} = \theta.$$

Then the leading eigenvalue is given as

$$(9.10) \qquad \lambda_* = \sigma_{\mathbf{f}} - \sigma_{\mathbf{s}} - \operatorname{sgn}(\cos(x_*))\sqrt{\sigma_{\mathbf{s}}^2 - \left(\tfrac{v_0 x_*}{2L}\right)^2},$$

and we have

$$\varphi(r, v) = \phi(r)\mathbb{1}_{\{v=v_0\}} + \phi(-r)\mathbb{1}_{\{v=-v_0\}}, \quad \tilde{\varphi}(r, v) = \phi(-r)\mathbb{1}_{\{v=v_0\}} + \phi(r)\mathbb{1}_{\{v=-v_0\}},$$

where

$$(9.11) \qquad \phi(r) = \frac{\sin\left(\tfrac{1}{2}x_*(1 - \tfrac{r}{L})\right)}{\sin(\tfrac{1}{2}x_*)}, \qquad r \in (-L, L).$$

Before turning to the proof of this proposition, there are a number of remarks we should make.

First, note that the condition $\theta = 1$ can be written $v_0/\sigma_{\mathsf{s}} = 2L$, and thus the different forms of the lead eigenfunction correspond intuitively to whether the average distance that a particle travels between scatterings is greater than, or less than, the width of the strip. Second, the symmetric form of φ above can be readily seen from the invariance of the system under the reflection $(r, v) \mapsto (-r, -v)$. Third, as seen in the previous discussion, the eigenfunctions φ and $\tilde{\varphi}$ are unique up to a multiplicative constant. Here, we fix this constant by requiring $\varphi(0, v_0) = \tilde{\varphi}(0, -v_0) = 1$.

9.4. Numerical experiments with the one-dimensional slab reactor. For this setting, we can easily develop numerical solutions for the NBP, NRW, and h-NRW, although we must decide which h function to work with in advance. As V consists of just two elements, the setting of Urts functions is easy to work with. For convenience, assume that $v_0/2L\sigma_{\mathsf{s}} < 1$; then we will work with an Urts function of the form

$$(9.12) \quad h_1(r, v_0) = \left(r + L + \frac{v_0}{\sigma_{\mathsf{s}}}\right) \wedge (L - r) \quad \text{and} \quad h_1(r, -v_0) = (r + L) \wedge \left(L - r + \frac{v_0}{\sigma_{\mathsf{s}}}\right)$$

for $-L < r < L$. In that case, from (2.7), we have that $\beta(r, v) = \sigma_{\mathsf{f}}$. Note, moreover, that, for example, when $(r, v) = (L, v_0)$,

$$\mathsf{L}h_1(L, v_0) = (\mathtt{A} - \beta)h_1(L, v_0) = -v_0 + \sigma_{\mathsf{s}}\left(h_1(L, -v_0) - h_1(L, v_0)\right) = 0.$$

Hence, we can easily verify that

$$\sup_{r \in D, v \in \{-v_0, v_0\}} \frac{\mathsf{L}h_1(r, v)}{h_1(r, v)} < \infty.$$

We will also consider the choices of the Urts functions proposed similar to those in Examples 8.3.

$$(9.13) \quad h_2(r, v_0) = L - r \quad \text{and} \quad h_3(r, v_0) = \left(r + L + \frac{v_0}{\sigma_{\mathsf{s}}}\right)(L - r),$$

and $h_i(r, -v_0) = h_i(-r, v_0)$ for $i = 2, 3$.

NBP estimators. Recall from Theorem 5.1 that the branching estimator $\Psi_k^{\mathtt{br}}[g](t, r, v)$ has different behaviors for supercritical, critical, and subcritical systems. While estimations improve as time grows in a supercritical situation, this is not the case for subcritical systems (Figure 5) since most genealogies have become extinct at large times. Moreover, simulations (Figure 6) suggest that in the supercritical case for fixed k the convergence rate is indeed $1/t$.

Next, we illustrate how our estimates of the simulation costs can be applied in this setting with two natural choices of (f, g): one being $f = \mathbf{0}$ and $g = \mathbf{1}$, which, as discussed previously, estimates the memory cost as the function $C[\mathbf{0}, \mathbf{1}](t)$ counts precisely the number of particles that have appeared before time t, the other chosen to be $f = \mathbf{1}$ and $g = \mathbf{0}$ for an estimate of the CPU time. Recall that in the supercritical case Lemma 5.2 predicts an exponential
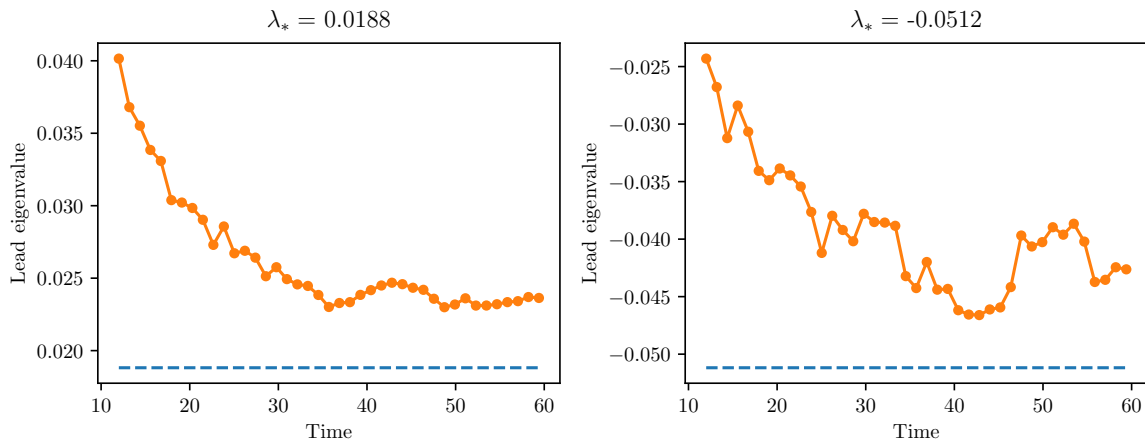
Figure 5. Supercritical vs. subcritical. *Estimates of λ_* from the branching estimator in a supercritical case (left) and a subcritical one (right).*
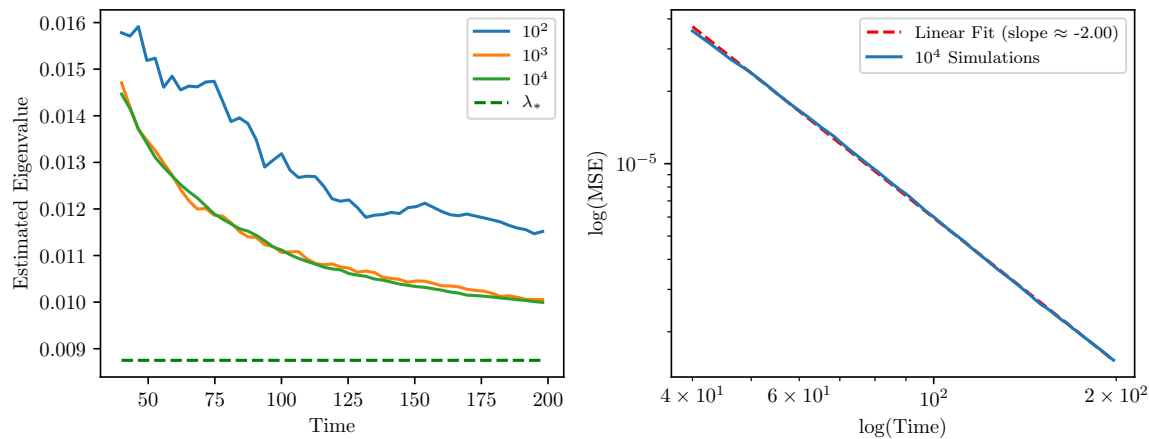

Figure 6. Convergence rate of branching estimators. *(Left) Estimates of λ_* from the branching estimator $\frac{1}{t}\Psi_k^{\mathrm{br}}[g](t,r,v)$ in a supercritical situation; the dashed line depicts the value of λ_*, and different lines correspond to experiments with different values of k. (Right) The function $t \mapsto (\frac{1}{t}\Psi_k^{\mathrm{br}}[g](t,r,v) - \lambda_*)^2$ in log-log scale; the slope is approximately -2.00.*

growth at rate λ_*. Note also that in the current setting, the various constants appearing in Lemma 5.2 can be easily computed from the solutions given in Proposition 9.1. See Figures 7 and 8 for the numerical results.

NRW and h-NRW estimators. Since the particles are typically killed on the boundary after some period of time, the basic NRW estimator $\Psi_k^{\mathrm{rw}}[g](t,r,v)$ performs worse than the branching estimator (Figure 9).

The h-NRW estimators depend on the choice of h. We see this by considering three cases based on (9.12) and (9.13). See Figure 10 for a plot of the h's and a comparison of the estimators they yield. The pictures there suggest that both h_1-NRW and h_3-NRW estimators
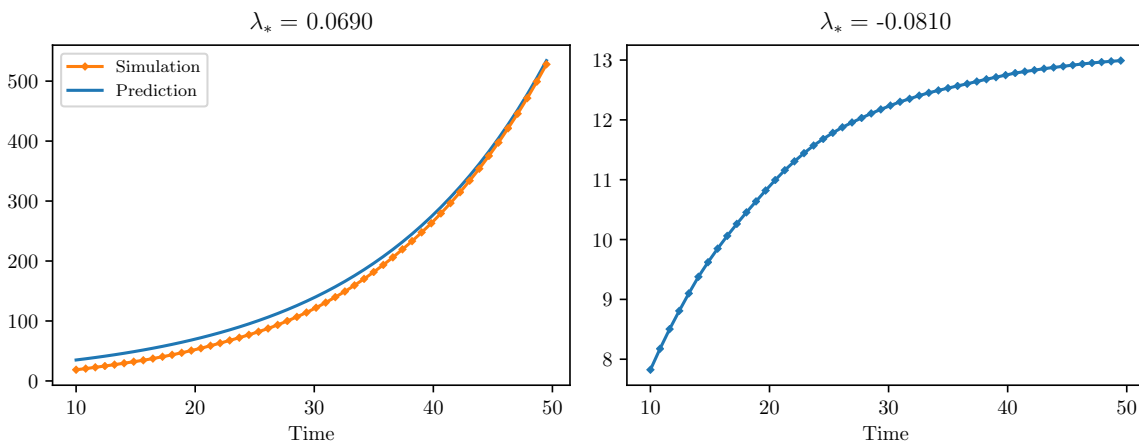
Figure 7. Supercritical vs. subcritical. *Mean memory costs in a supercritical case (left) and a subcritical case (right) for the NBP. In the supercritical case, this is compared to the prediction derived from* (5.6).
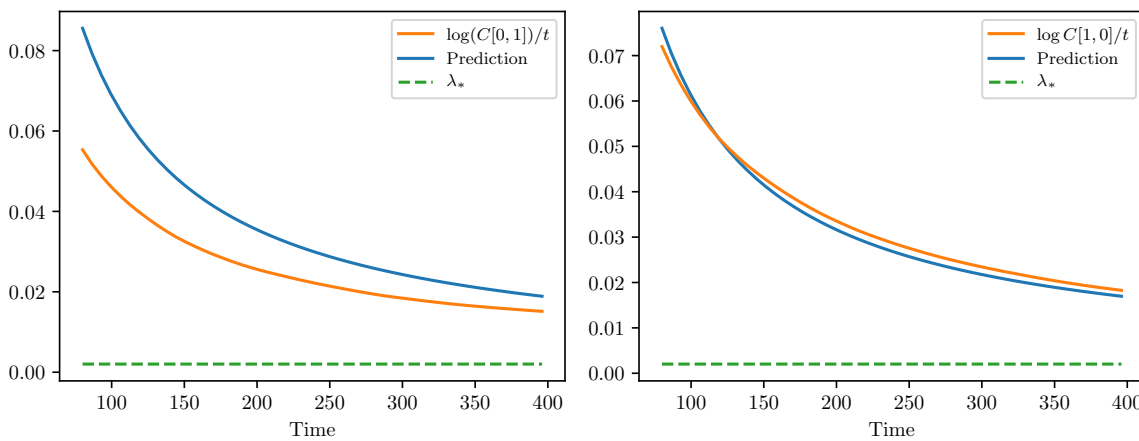

Figure 8. *(Left) The orange curve is the plot of* $t \mapsto \frac{1}{t} \log C[\mathbf{0}, \mathbf{1}](t)$ *averaged over* 500 *simulations of the NBP. The blue line depicts the predicted growth rate given on the right-hand side of* (5.6), *and the green line indicates the value of* λ_*. *(Right) The orange curve is the plot of* $t \mapsto \frac{1}{t} \log C[\mathbf{1}, \mathbf{0}](t)$ *averaged over* 500 *simulations of the NBP. The blue line depicts the predicted growth rate given on the RHS of* (5.6), *and the green line indicates the value of* λ_*.

outperformed the NBP estimator in this example. Note also, based on the discussion in the previous section, that the "poor" choice of h, *viz.* h_2, systematically underestimates λ_*.

9.5. A two-dimensional reactor with four rods. We compare the results above with a two-dimensional particle simulation. In this case, we model our environment as a square reactor with 4 rods, where most branching happens, and with particles killed on exit from the square. We assume all scattering events are uniform, and the particles have constant speed.

Numerically, the picture is similar to the one-dimensional case. The basic NRW estimator suffers from the rapid loss of the particles (Figure 9). For the h-NRW, we use the choice of h given in (8.7). Estimates over time for the eigenvalue by the branching estimator and the h-NRW estimator can be found in Figure 11.
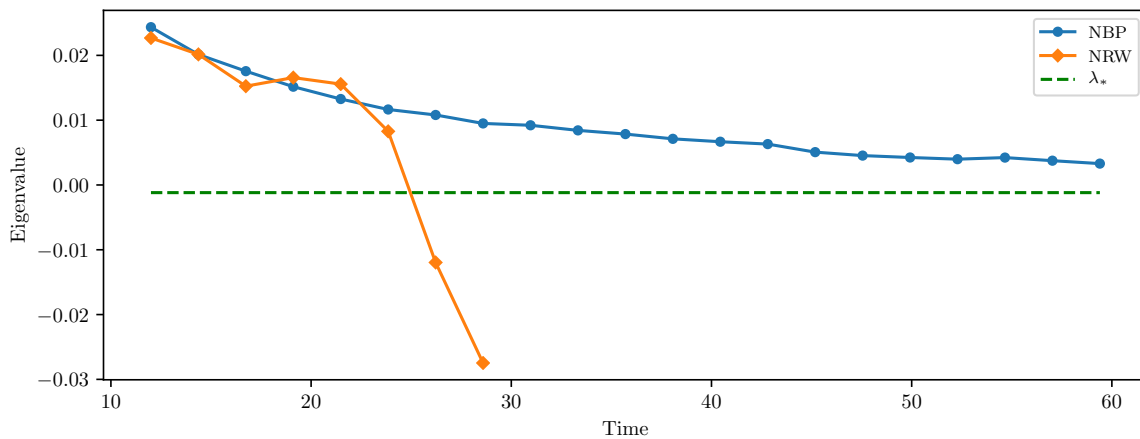
Figure 9. NBP vs. NRW. *Estimates of λ_* for the NBP and NRW estimators in a supercritical case. In the NRW case, the plot stops after all particles are killed. The instability is due to the very small number of surviving particles at large times. The true eigenvalue here is $\lambda_* \approx -0.00118$.*
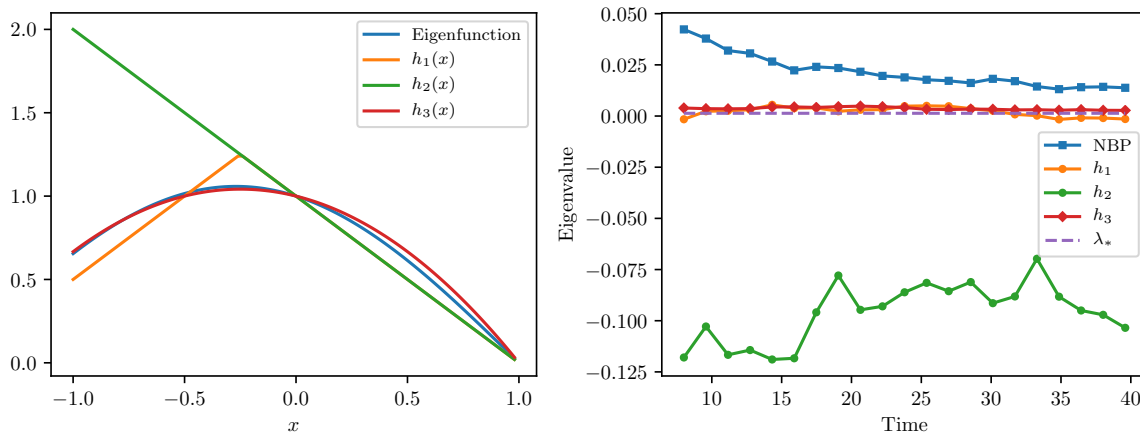
Figure 10. h-NRW estimators. *(Left) Plots of three different choices of h compared to the true eigenfunction (in blue). (Right) Comparison of these three h-NRW eigenvalue estimators, together with the NBP estimator ($k = 1000$ simulations).*

In Figure 11 we see the advantages and disadvantages of the h-transform method. In the simplest case (left), where particles are simulated until death, the NRW suffers from extinction of its particles, even though a large number of particles (50,000) are initially simulated. As a result, the estimate of λ_* is very unreliable, and there is little hope in getting the particles close to equilibrium before most particles die out. The h-NRW case does much better on this count, with particles surviving until large times; however, the particles in this case suffer from weight degeneracy, where a few particles have large weight (the exponential, product, and indicator terms in (7.11)), and many particles have comparatively small weight. As a result, the final computations of the average will be dominated by the small number of particles with large weight, leading to wasted computational effort.
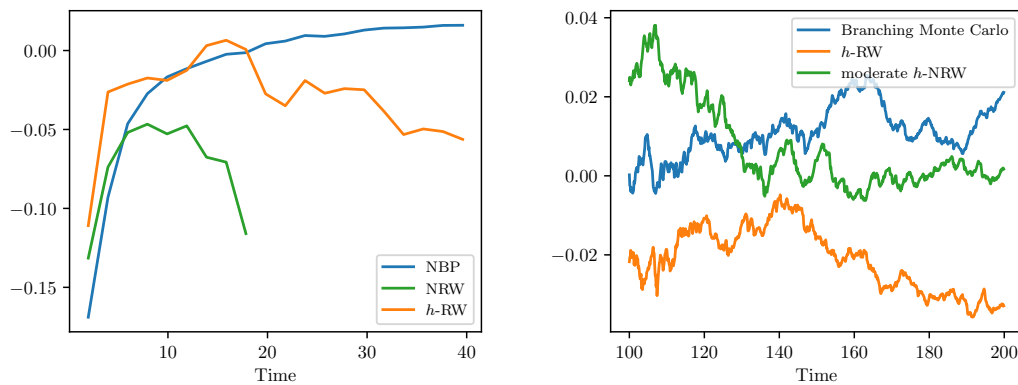
Figure 11. *h-NRW estimators for 2D reactor. With the choice of h given in* (8.7). *Here we choose* $\eta = 0.02, C = 0.1$. *(Left) The estimates for the NBP, NRW, and h-NRW cases. The particles in the NRW (started with* 50,000 *particles) all die out by time* 20. *The h-transform estimate is based on* 200 *particles, and the NRW estimate is for* 400 *particles. The drop-off in the estimates of the h-transform method appear to be due to weight degeneracy of the particles. (Right) The eigenvalue estimated using a particle filter approach. Particles are resampled based on the weight of current particles to minimize the effects of weight degeneracy. The "moderate h-NRW" uses a mild version of the h-transform which is closer to the NRW.*

To mitigate the weight degeneracy of the particles, we implement a particle filter algorithm. Roughly speaking, this means that at fixed times, we look at our collection of particles and randomly choose to discard/keep particles based on the weight of individual particles, for example, by choosing particles in the next generation by independently selecting particles from the current population proportional to their current weights. In our implementation, we use a simple multinomial resampling scheme. It is known that more sophisticated resampling schemes can improve the performance of the filter [19], although we do not expect substantial benefits in this setting. We also resample our particles at fixed time-steps. More sophisticated algorithms might choose only to resample when a simple measure of the weight diversity, such as the *effective sample, size* drops below a chosen threshold. Since we are not trying to measure directly the algorithmic performance in this paper, we have not tried to complicate our implementation of the particle filter in this paper.

Particle filtering algorithms have been well studied; see, for example, [18] or [1]. In this case, the *h*-transform methods appear to be slightly more stable than the branching process approximation, although it is hard to judge the accuracy here due to the lack of an analytical solution. In addition, in our implementation of this method, the *h*-transform approach is substantially slower, due to the complexity of simulating paths from the conditioned process, compared to the NBP. However in industrial implementations, the complexity of the NBP will increase as the environment gets more sophisticated, and so conditioning may become more competitive with the NBP approach, particularly when combined with smarter ways to calibrate the *h* function. We leave details of these improvements as the subject of future work.

Appendix A. Monte Carlo algorithms. In this appendix, we highlight briefly the approach to simulating an NRW, an NBP, and an NRW with importance sampling, which themselves feed into the basic Monte Carlo algorithms in sections 4 and 7. We assume that the basic data of D, V and $\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}}$ are all given.

Algorithm A.1 (simulation of an $\alpha\pi$-NRW).

Given an initial configuration (r, v), an initial time $\mathtt{t}_0$, and a time horizon t:

Step 1: Set $n = 0$, $(\mathtt{r}_0, \mathtt{v}_0) = (r, v) \in D \times V$.

Step 2: Given the value of $\mathtt{t}_n$, $\mathtt{r}_n$, $\mathtt{v}_n$, sample δ_{n+1} with law $\mu^{\mathtt{s}}_{\mathtt{r}_n, \mathtt{v}_n}$ and set

$$\mathtt{t}_{n+1} = \mathtt{t}_n + \delta_{n+1}, \quad \mathtt{r}_{n+1} = \mathtt{r}_n + \mathtt{v}_n \delta_{n+1}.$$

Step 3: Check if $\mathtt{r}_{n+1} \in D$ and $\mathtt{t}_{n+1} < t$. If true, sample $\mathtt{v}_{n+1}$ with law $\eta^{\mathtt{s}}_{\mathtt{r}_{n+1}, \mathtt{v}_n}$, increase n by one, and go to Step 2. If false, set

$$\mathtt{t}_{\text{end}} = \min\{\mathtt{t}_n + \inf\{s > 0 : \mathtt{r}_n + \mathtt{v}_n s \notin D\}, t\}.$$

Step 4: Return $((\mathtt{t}_k, \mathtt{r}_k, \mathtt{v}_k), 0 \le k \le n)$ and $\mathtt{t}_{\text{end}}$.

A.1. Simulating $\alpha\pi$-NRW. We start by defining the distributions

$$\mu^{\mathtt{s}}_{r,v}(u, \infty) := e^{-\int_0^u \alpha(r + vs, v)\,ds}, \qquad u \ge 0, r \in D, v \in V,$$

and

$$\eta^{\mathtt{s}}_{r,v}(dv') = \pi(r, v, v')dv', \qquad r \in D, v \in V, v' \in V.$$

The algorithm below for generating an NRW will take as input the starting position and velocity, (r, v), and the terminal time t. For later use, we will also allow the particle to start at a positive time $\mathtt{t}_0 \in [0, t)$.

We note that there are further algorithmic simplifications within the above description. For example, sampling from the distribution $\mu^{\mathtt{s}}_{r,v}$ can be done as follows. Suppose that $\mathbf{e}$ is a unit mean exponential random variable. For each $r \in D$ and $v \in V$, the law $\mu_{r,v}(\cdot)$ agrees with that of $\inf\{t > 0 : \int_0^t \beta(r + vs, v)ds > \mathbf{e}\}$. In the case that β is bounded above by a constant, this is further simplified through the use of a simple rejection sampling algorithm.

The above algorithm returns a sequence of variables $\mathbf{z} := \{((\mathtt{t}_k, \mathtt{r}_k, \mathtt{v}_k), 0 \le k \le n), \mathtt{t}_{\text{end}}\}$, for some n, and a final time $\mathtt{t}_{\text{end}} \le t$. Together these fully determine the trajectory, say $((\mathtt{r}^{\mathbf{z}}_s, \mathtt{v}^{\mathbf{z}}_s), s < \mathtt{t}_{\text{end}})$, of an NRW, in the following way:

$$\text{(A.1)} \qquad \mathtt{r}^{\mathbf{z}}_s = \mathtt{r}_k + \mathtt{v}_k(s - \mathtt{t}_k) \quad \text{and} \quad \mathtt{v}^{\mathbf{z}}_s = \mathtt{v}_k \quad \text{if} \quad \mathtt{t}_k \le s < \mathtt{t}_{k+1} \wedge \mathtt{t}_{\text{end}},$$

where we take $\mathtt{t}_{n+1} = \infty$. By simulating many particles in this manner, the value of the semigroup may be estimated via (4.2).

A.2. Simulating $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \mathcal{P})$-NBP when only $\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}}$ is given. Building on section A.1, we have the algorithm below, which generates the branching particle system. We need the following notation. Let $\mathbf{z}$ be a trajectory as defined above. Then we define a measure on $[\mathtt{t}^{\mathbf{z}}_0, \infty) \cup \{\infty\}$ by

$$\mu^{\mathtt{f}}_{\mathbf{z}}((u, \infty) \cup \{\infty\}) := e^{-\int_{\mathtt{t}^{\mathbf{z}}_0}^{u \wedge \mathtt{t}^{\mathbf{z}}_{\text{end}}} \sigma_{\mathtt{f}}(\mathtt{r}^{\mathbf{z}}_s, \mathtt{v}^{\mathbf{z}}_s)\,ds}, \qquad u \ge \mathtt{t}^{\mathbf{z}}_0,$$

and a measure on V by

$$\eta_{\mathbf{z},t}^{\mathbf{f}}(\mathrm{d}v') = \frac{\pi_{\mathbf{f}}(\mathbf{r}_t^{\mathbf{z}}, \mathbf{v}_t^{\mathbf{z}}, v')\mathrm{d}v'}{\pi_{\mathbf{f}}(\mathbf{r}_t^{\mathbf{z}}, \mathbf{v}_t^{\mathbf{z}}, V)}, \qquad v' \in V.$$

In order to simulate an NBP with given data $\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \pi_{\mathbf{f}}$, we need to supply additional information, as the family of densities $\pi_{\mathbf{f}}(r, v, \cdot)$, $r \in D, v \in V$, only tells us the mean behavior of the point process of outgoing fission velocities, whose probabilities were denoted by $\mathcal{P}_{(r,v)}$, $r \in D, v \in V$. Moreover, there is no unique way to choose $\mathcal{P}_{(r,v)}$, $r \in D, v \in V$, given $\pi_{\mathbf{f}}(r, v, \cdot)$, $r \in D, v \in V$.

This does not present a problem, however; on the contrary it presents an opportunity. As all of our estimators that are based on the NBP are built around the notion of mean growth, we are at liberty to choose a convenient $(\mathcal{P}_{(r,v)}, r \in D, v \in V)$, and, as we shall see below, an obvious choice is that of a Poisson random field with intensity density given by $\pi_{\mathbf{f}}(r, v, \cdot)$, $r \in D, v \in V$. As a small remark, this choice is unrealistic as a physical model given the assumption (H2); however, for the synthetic purposes of Monte Carlo simulation, it is very convenient both practically and mathematically.

Below, we give the algorithm to produce the paths of the NBP. Note that, unlike the NRW, the output of the NBP is a random number of trajectories that may have birth times $b^i = \mathbf{t}_0^i$ and death times $\mathbf{t}_{\mathrm{end}}^i$ which may be strictly positive, and strictly before the end time t. The object returned will therefore be a set of trajectories (via (A.1)) of the form $\{((\mathbf{t}_k^i, \mathbf{r}_k^i, \mathbf{v}_k^i), 0 \leq k \leq n^i), \mathbf{t}_{\mathrm{end}}^i\}_{i=1}^N$, where N is the (random) total number of particles in the branching process.

Algorithm A.2 (simulation of a $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \pi_{\mathbf{f}})$-NBP).

Given an initial configuration $(r_1, v_1), \ldots (r_\ell, v_\ell)$ and a time horizon t:

Step 1: Set $\ell^0 = \ell$, $b_i = 0, i = 1, \ldots, \ell^0$, $\mathcal{X}_0 = \{(r_1, v_1, b_1), \ldots, (r_\ell, v_\ell, b_\ell)\}$, and $n = 0$.

Step 2: For each $i \in \{1, \ldots, \ell^n\}$ and corresponding $(r_i, v_i, b_i) \in \mathcal{X}_n$ such that $b_i < t$, run Algorithm A.1 to produce a $\sigma_{\mathbf{s}}\pi_{\mathbf{s}}$-NRW with initial configuration (r_i, v_i), birth time b_i, and terminal time t. Denote the resulting NRW by
$\mathbf{z}^{i,n} := \{((\mathbf{r}_k^{i,n}, \mathbf{t}_k^{i,n}, \mathbf{v}_k^{i,n}), 0 \leq k \leq m^{i,n}), \mathbf{t}_{\mathrm{end}}^{i,n}\}$.

Step 3: For each $i \in \{1, \ldots, \ell^n\}$ and corresponding $(r_i, v_i, b_i) \in \mathcal{X}_n$ such that $b_i < t$, sample $\gamma^{i,n}$ with law $\mu_{\mathbf{z}^{i,n}}^{\mathbf{f}}$ and then, if $\gamma^{i,n} \leq \mathbf{t}_{\mathrm{end}}^{i,n}$:

1. Sample $\xi^{i,n}$ independent variables $u_1^{i,n}, \ldots, u_{\xi^{i,n}}^{i,n}$ with probability density $\eta_{\mathbf{z}^{i,n}}^{\mathbf{f}}$, where $\xi^{i,n}$ is Poisson distributed with parameter $\pi_{\mathbf{f}}(\mathbf{r}_{\gamma^{i,n}}^{i,n}, \mathbf{v}_{\gamma^{i,n}}^{i,n}, V)$; and

2. Trim the path $\mathbf{z}^{i,n}$ at time $\gamma^{i,n}$, that is, set $\tilde{m}^{i,n}$ to be the largest integer $k \leq m^{i,n}$ such that $\mathbf{t}_k^{i,n} \leq \gamma^{i,n}$, and redefine $\mathbf{z}^{i,n} := \{((\mathbf{r}_k^{i,n}, \mathbf{t}_k^{i,n}, \mathbf{v}_k^{i,n}), 0 \leq k \leq \tilde{m}^{i,n}), \gamma^{i,n}\}$.

Step 4: Let $\mathcal{X}_{n+1}$ be the set $\{(\mathbf{r}_{\gamma^{i,n}}^{i,n}, u_j^{i,n}, \gamma^{i,n}); 1 \leq i \leq \ell^n, 1 \leq j \leq \xi^{i,n}, \mathbf{t}_0^{i,n} < t\}$. Set $\ell^{n+1} := |\mathcal{X}_{n+1}|$.

Step 5: If $\ell^{n+1} > 0$, go to Step 2. Otherwise stop and return $\mathcal{X} = \{\mathbf{z}^{i,n}; 1 \leq i \leq \ell^n, n \in \mathbb{N}\}$.

In this algorithm, we then estimate the semigroup using (4.1), so we have (in the notation of the algorithm above)

$$\Psi_k[g](t, r, v) = \frac{1}{k} \sum_{\mathbf{z} \in \mathcal{X}} g(\mathbf{r}_t^{\mathbf{z}}) \mathbf{1}_{(\mathbf{t}_{\mathrm{end}}^{\mathbf{z}} \geq t)}.$$

Let us make some practical notes on the simulation of a $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}})$-NBP, pertaining in particular to how such simulations are performed in industrial software. The geometric structure of a virtual reactor will be designed on a special CAD tool. After that, according to the physical properties of the materials used, the values of the four cross-sections $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}})$ are mapped from data libraries into the different geometrical regions. This generates numerically the piecewise constant functions $(\sigma_{\mathtt{s}}, \pi_{\mathtt{s}}, \sigma_{\mathtt{f}}, \pi_{\mathtt{f}})$ across the entire space-velocity domain $D \times V$. The construction of this numerical representation of the cross-sections in a virtual nuclear reactor can be stored in a file whose size[10] is as little as 150MB. Thereafter, the NBP simulation calls upon the values in this file as it samples from, e.g., $\mu_{r,v}^{\mathtt{f}}$ and $\eta_{r,v}^{\mathtt{f}}$.

A.3. Simulating $\alpha\pi$-NRW Monte Carlo with importance sampling. In conclusion, a second proposed efficiency to the Monte Carlo simulation of solutions to the NTE is to make an educated guess at the shape of φ and the function h and then to adapt Algorithm A.1 as follows.

Algorithm A.3.

Run Step 1–4 *of Algorithm* A.1 *replacing the quantities α, β, and π by*

$$\alpha^h(r, v) = \frac{\alpha(r, v) \int_V h(r, v') \pi(r, v, v') \, \mathrm{d}v'}{h(r, v)}, \qquad \beta^h(r, v) = \frac{\mathtt{J}h(r, v)}{h(r, v)} + \beta(r, v),$$

and

$$\pi^h(r, v, v') = \frac{h(r, v')}{\int_V h(r, v') \pi(r, v, v') \, \mathrm{d}v'} \pi(r, v, v'),$$

noting that $\mathtt{t}^{\mathrm{end}} = t$ *always.*

In this case, the Monte Carlo estimate of the operator $\Psi_k[g]$ is then given by (8.1).

Appendix B. A many-to-two lemma for NBP. Recall the notation of (2.1), (2.2) and define, for $r \in D$, $v \in V$, $f, g \in L_\infty^+(D \times V)$,

(B.1) $$\eta_{\mathtt{f}}[f, g](r, v) = \sigma_{\mathtt{f}}(r, v) \mathcal{E}_{(r,v)} \left[\sum_{i,j=1, i \neq j}^N f(v_i) g(v_j) \right].$$

---

[10]This is information which has been shared with us by our industrial partner at the ANSWERS group at Jacobs.

As an abuse of notation, we will also write, for $r \in D$, $v \in V$, $f \in L_\infty(D \times V)$,

$$\eta_{\mathtt{f}}[f](r,v) := \eta_{\mathtt{f}}[f,f](r,v) = \sigma_{\mathtt{f}}(r,v)\mathcal{E}_{(r,v)}\left[\sum_{i,j=1,i \neq j}^{N} f(v_i)f(v_j)\right].$$

We have the following two-point correlation formula (also called the many-to-two formula; cf [27]).

Lemma B.1 (many-to-two). *Suppose that $f,g \in L_\infty^+(D \times V)$. Then*

$$(B.2) \qquad \mathbb{E}_{\delta(r,v)}\Big[\langle f, X_t\rangle\langle g, X_t\rangle\Big] = \psi_t[fg](r,v) + \int_0^t \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[f], \psi_{t-s}[g]]\Big](r,v)\mathrm{d}s.$$

Proof. The proof follows the arguments in Harris and Roberts [27], which can be easily adapted to the current context.

As a first step, we note that for $h \in L_\infty^+(D \times V)$ and $H : \mathbb{R}_+ \times D \times V \to \mathbb{R}_+$ continuous and bounded, the function

$$u_t(r,v) := \mathbf{E}_{(r,v)}\left[\mathrm{e}^{\int_0^t \beta(R_s, \Upsilon_s)\mathrm{d}s}h(R_t, \Upsilon_t)\mathbf{1}_{(t<\tau^D)}\right] + \mathbf{E}_{(r,v)}\left[\int_0^{t \wedge \tau^D} H(s, R_s, \Upsilon_s)\mathrm{e}^{\int_0^s \beta(R_u, \Upsilon_u)\mathrm{d}u}\mathrm{d}s\right]$$

$$(B.3) \qquad = \psi_t[h](r,v) + \mathbf{E}_{(r,v)}\left[\int_0^{t \wedge \tau^D} H(s, R_s, \Upsilon_s)\mathrm{e}^{\int_0^s \beta(R_u, \Upsilon_u)\mathrm{d}u}\mathrm{d}s\right],$$

for $t \geq 0$, $r \in D$, and $v \in V$, uniquely solves the integral equation

$$(B.4) \qquad u_t(r,v) = \mathtt{U}_t[h](r,v) + \int_0^t \mathtt{U}_s\Big[H(s,\cdot) + (\mathtt{S}+\mathtt{F})u_{t-s}\Big](r,v)\mathrm{d}s.$$

We only sketch the proof of this fact for the sake of brevity. To see why this is the unique solution, it suffices to return to the proof of, e.g., Lemma 6.1 in [13] and note that the approach given there works equally well here. That is to say, we first condition the right-hand side of (B.3) on the first fission or scattering event, whichever comes first, to generate a recursion in u. This can then be manipulated further with the help of Lemma 1.2 in of Chapter 4 [21] to obtain (B.4). Finally uniqueness of the latter follows by a standard argument appealing to Grönwall's lemma.

To complete the proof of (B.2), it suffices to consider the case $f = g$, as the general form will follow from the polarization

$$\mathbb{E}_{\delta(r,v)}\Big[2\langle f, X_t\rangle\langle g, X_t\rangle\Big] = \mathbb{E}_{\delta(r,v)}\Big[\langle f+g, X_t\rangle^2\Big] - \mathbb{E}_{\delta(r,v)}\Big[\langle f, X_t\rangle^2\Big] - \mathbb{E}_{\delta(r,v)}\Big[\langle g, X_t\rangle^2\Big].$$

To this end, denote by $w_t(r,v) = \mathbb{E}_{\delta(r,v)}[\langle g, X_t\rangle^2]$ for $t \geq 0$, $r \in D$, $v \in V$. For convenience, write $\Pi_t(r,v) = \exp(-\int_0^t \sigma(r+\ell v, v)\mathrm{d}\ell)$ for $r \in D$, $v \in V$. Once again, by conditioning on the first scattering or fission event and then applying Lemma 1.2 of [21], we get

$$w_t(r,v) = \Pi_t(r,v)\mathtt{U}_t[g^2](r,v) + \int_0^t \mathtt{U}_s\left[\sigma_{\mathtt{s}}\Pi_s\int_V w_{t-s}(r,v')\pi_{\mathtt{s}}(\cdot,\cdot,v')\mathrm{d}v'\right](r,v)\mathrm{d}s$$

$$+ \int_0^t \mathtt{U}_s\left[\sigma_{\mathtt{f}}\Pi_s\mathcal{E}_{(r,v)}\left[\sum_{i,j=1,i\neq j}^N \psi_{t-s}[g](\cdot,v_i)\psi_{t-s}[g](\cdot,v_j) + \sum_{i=1}^N w_{t-s}[g](\cdot,v_i)\right]\right](r,v)\mathrm{d}s$$

$$= \mathtt{U}_t[g^2](r,v) + \int_0^t \mathtt{U}_s[(\mathtt{S}+\mathtt{F})w_{t-s}](r,v)\mathrm{d}s + \int_0^t \mathtt{U}_s\left[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\right](r,v)\mathrm{d}s.$$

Using the representation of the solution to (B.3) with $h = g^2$ and $H(s,\cdot,\cdot) = \eta_{\mathtt{f}}[\psi_{t-s}[g]]$, we get (B.2). ∎

Appendix C. NBP Monte Carlo convergence.

Proof of Theorem 5.1. The estimator $\Psi_k^{\mathtt{br}}$ as defined in (4.1) is unbiased. Essentially the proof is based on a study of its variance, which amounts to studying the variance of the branching system, since

$$(\text{C.1}) \qquad \mathbb{E}_{\delta(r,v)}\left[\left(\Psi_k^{\mathtt{br}}[g](t,r,v) - \psi_t[g](r,v)\right)^2\right] = \frac{1}{k}\left(\mathbb{E}_{\delta(r,v)}[\langle g, X_t\rangle^2] - \psi_t[g](r,v)^2\right).$$

We shall rely upon the many-to-two formula in Lemma B.1 to compute the above. Accordingly, we now claim the following as an intermediate step.

Lemma C.1. *We have for all $t \geq 0$, $k \in \mathbb{N}$,*

$$(\text{C.2}) \qquad \mathbb{E}_{\delta(r,v)}\left[\left(\Psi_k^{\mathtt{br}}[g](t,r,v) - \psi_t[g](r,v)\right)^2\right]$$

$$= \frac{1}{k}\left\{\psi_t[g^2](r,v) + \int_0^t \psi_s\left[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\right](r,v)\mathrm{d}s - \psi_t[g](r,v)^2\right\}.$$

Subsequently, it has the following asymptotics as $t \to \infty$:

(i) *If $\lambda_* = 0$, then for fixed k,*

$$\lim_{t\to\infty}\frac{k}{t}\cdot\mathbb{E}_{\delta(r,v)}\left[\left(\Psi_k^{\mathtt{br}}[g](t,r,v) - \psi_t[g](r,v)\right)^2\right] = \langle\eta_{\mathtt{f}}[\varphi],\tilde\varphi\rangle\langle g,\tilde\varphi\rangle^2\varphi(r,v) := C_{[1]}(g,r,v).$$

(ii) *If $\lambda_* > 0$, then for fixed k,*

$$\lim_{t\to\infty}e^{-2\lambda_* t}k\cdot\mathbb{E}_{\delta(r,v)}\left[\left(\Psi_k^{\mathtt{br}}[g](t,r,v) - \psi_t[g](r,v)\right)^2\right]$$

$$= \langle g,\tilde\varphi\rangle^2\left(\int_0^\infty e^{-2\lambda_* s}\psi_s[\eta_{\mathtt{f}}[\varphi]](r,v)\mathrm{d}s - \varphi(r,v)^2\right) := C_{[2]}(g,r,v) \in [0,\infty).$$

(iii) *If $\lambda_* < 0$, then for fixed k,*

$$\lim_{t\to\infty}e^{-\lambda_* t}k\cdot\mathbb{E}_{\delta(r,v)}\left[\left(\Psi_k^{\mathtt{br}}[g](t,r,v) - \psi_t[g](r,v)\right)^2\right]$$

$$= \varphi(r,v)\left\{\langle g^2,\tilde\varphi\rangle + \int_0^\infty e^{-\lambda_* s}\langle\tilde\varphi,\eta_{\mathtt{f}}[\psi_s[g]]\rangle\mathrm{d}s\right\} := C_{[3]}(g,r,v) \in [0,\infty).$$

Proof. The display (C.2) is an easy consequence of (B.2). We turn to the asymptotics. Appealing to (3.2), we have for any $f \in L_\infty^+(D \times V)$ and $\delta > 0$ that there exists some constant $K \in (0, \infty)$ and some $t_0 = t_0(\delta)$ such that

(C.3) $\quad \sup_{t \geq 0} \|e^{-\lambda_* t} \psi_t[f]\|_\infty \leq K\|f\|_\infty \quad$ and $\quad \|e^{-\lambda_* t} \psi_t[f] - \langle f, \tilde{\varphi} \rangle \varphi\|_\infty \leq \delta\|f\|_\infty \quad$ for all $t \geq t_0$.

On the other hand, it follows from (B.1) that $\eta_{\mathtt{f}}$ is a symmetric bilinear form, and hence

(C.4) $\quad |\eta_{\mathtt{f}}[f](r, \upsilon) - \eta_{\mathtt{f}}[h](r, \upsilon)| = |\eta_{\mathtt{f}}[f - h, f + h](r, \upsilon)| \leq C\|f - h\|_\infty (\|f\|_\infty + \|h\|_\infty),$

where in the last inequality we have used the fact that $\|\sigma_{\mathtt{f}}\|_\infty < \infty$ and the offspring number is also uniformly bounded by some constant B, so that above, $C := \|\sigma_{\mathtt{f}}\|_\infty B < \infty$. In particular, taking $h = 0$ yields

(C.5) $$\|\eta_{\mathtt{f}}[f]\|_\infty \leq C\|f\|_\infty^2.$$

Also, we clearly have

(C.6) $$f, g \in L_\infty, f \leq g \quad \Rightarrow \quad \psi_t[f] \leq \psi_t[g], \quad t \geq 0.$$

Let us look at the supercritical case, i.e., $\lambda_* > 0$. The arguments consist in extracting the dominant growth rate from (C.2) as $t \to \infty$. To that end, we first split the integral in (C.2) into two parts:

$$\int_0^t \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\Big](r, \upsilon)\mathrm{d}s = \int_{t-t_0}^t \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\Big](r, \upsilon)\mathrm{d}s + \int_0^{t-t_0} \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\Big](r, \upsilon)\mathrm{d}s.$$

Note that the first term on the right-hand side is of order $o(e^{2\lambda_* t})$. Indeed,

$$\int_{t-t_0}^t \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\Big](r, \upsilon)\mathrm{d}s \leq \int_{t-t_0}^t \psi_s\Big[\eta_{\mathtt{f}}[K\|g\|_\infty e^{\lambda_*(t-s)}]\Big](r, \upsilon)\mathrm{d}s \qquad \text{by (C.3)}$$

$$\leq \int_{t-t_0}^t \psi_s\Big[CK^2\|g\|_\infty^2 e^{2\lambda_*(t-s)}\Big](r, \upsilon)\mathrm{d}s \qquad \text{by (C.5)}$$

$$\leq CK^3\|g\|_\infty^2 e^{2\lambda_* t} \int_{t-t_0}^t e^{-\lambda_* s}\mathrm{d}s \qquad \text{by (C.6) and (C.3)}$$

$$= o(e^{2\lambda_* t}), \qquad \text{as } t \to \infty.$$

On the other hand, for the second term, we have

$$\Big| \int_0^{t-t_0} \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\Big](r, \upsilon)\mathrm{d}s - \langle g, \tilde{\varphi} \rangle^2 \int_0^{t-t_0} e^{2\lambda_*(t-s)} \psi_s\Big[\eta_{\mathtt{f}}[\varphi]\Big](r, \upsilon)\mathrm{d}s \Big|$$

$$\leq \int_0^{t-t_0} \psi_s\Big[\Big|\eta_{\mathtt{f}}[\psi_{t-s}[g]] - \eta_{\mathtt{f}}[\langle g, \tilde{\varphi} \rangle e^{\lambda_*(t-s)} \varphi]\Big|\Big](r, \upsilon)\mathrm{d}s$$

$$\leq \int_0^{t-t_0} \psi_s\Big[e^{2\lambda_*(t-s)} \delta CK\|g\|_\infty (\|g\|_\infty + |\langle g, \tilde{\varphi} \rangle| \|\varphi\|_\infty)\Big]\mathrm{d}s \quad \text{by (C.3), (C.4)}$$

$$\leq \delta CK\|g\|_\infty (\delta\|g\|_\infty + 2|\langle g, \tilde{\varphi} \rangle| \|\varphi\|_\infty) e^{2\lambda_* t} \int_0^{t-t_0} e^{-\lambda_* s}\mathrm{d}s \quad \text{by (C.3)}$$

$$= O(\delta e^{2\lambda_* t}), \qquad \text{as } t \to \infty.$$

Since δ is arbitrary, when combined with the fact that the integral

$$\int_0^\infty e^{-2\lambda_* s} \psi_s[\eta_{\mathtt{f}}[\varphi]](r,\upsilon)\mathrm{d}s \leq CK\|\varphi\|_\infty^2 \int_0^\infty e^{-\lambda_* s}\mathrm{d}s < \infty,$$

the above implies

$$\int_0^{t-t_0} \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s \overset{t\to\infty}{\sim} \langle g, \tilde\varphi\rangle^2 e^{2\lambda_* t}\int_0^\infty e^{-2\lambda_* s}\psi_s[\eta_{\mathtt{f}}[\varphi]](r,\upsilon)\mathrm{d}s.$$

The asymptotics in the supercritical case then easily follows. Next, we consider the subcritical case, i.e., $\lambda_* < 0$. We start with a change of variable:

$$\int_0^t \psi_s\Big[\eta_{\mathtt{f}}[\psi_{t-s}[g]]\Big](r,\upsilon)\mathrm{d}s = \int_0^t \psi_{t-s}\Big[\eta_{\mathtt{f}}[\psi_s[g]]\Big](r,\upsilon)\mathrm{d}s.$$

Note that by (C.3) and (C.5),

$$\int_{t-t_0}^t \psi_{t-s}\Big[\eta_{\mathtt{f}}[\psi_s[g]]\Big](r,\upsilon)\mathrm{d}s \leq CK^2\|g\|_\infty^2 \int_{t-t_0}^t e^{2\lambda_* s + \lambda_*(t-s)}\mathrm{d}s = o(e^{\lambda_* t}), \quad \text{as } t\to\infty,$$

since $\lambda_* < 0$. On the other hand, applying (C.3) to ψ_{t-s} and $\eta_{\mathtt{f}}[\psi_s[g]]$, noting the latter is bounded by $CK^2 e^{2\lambda_* s}\|g\|_\infty^2$ as consequence of (C.5) and (C.6), we get

$$\left|\int_0^{t-t_0} \psi_{t-s}\Big[\eta_{\mathtt{f}}[\psi_s[g]]\Big](r,\upsilon)\mathrm{d}s - e^{\lambda_* t}\varphi(r,\upsilon)\int_0^{t-t_0} e^{-\lambda_* s}\langle\tilde\varphi, \eta_{\mathtt{f}}[\psi_s[g]]\rangle\mathrm{d}s\right|$$

$$\leq \int_0^{t-t_0} \delta e^{\lambda_*(t-s)}\|\eta_{\mathtt{f}}[\psi_s[g]]\|_\infty\mathrm{d}s$$

$$\leq \int_0^{t-t_0} \delta e^{\lambda_*(t+s)}CK^2\|g\|_\infty^2\mathrm{d}s = O(\delta e^{\lambda_* t}), \quad \text{as } t\to\infty.$$

Arguing as in the previous case, we conclude that

$$\int_0^t \psi_{t-s}\Big[\eta_{\mathtt{f}}[\psi_s[g]]\Big](r,\upsilon)\mathrm{d}s \overset{t\to\infty}{\sim} e^{\lambda_* t}\varphi(r,\upsilon)\int_0^\infty e^{-\lambda_* s}\langle\tilde\varphi, \eta_{\mathtt{f}}[\psi_s[g]]\rangle\mathrm{d}s,$$

which in turn implies the result in the subcritical case. The proof in the critical case follows from similar arguments and is therefore omitted. ∎

We begin by noting the following estimates, which follow from the concavity of the function $h_t(x) := x^{1/t}$. Fix $x_0 \in (0,\infty)$, and consider $x \geq 0$. Then

$$0 \leq h_t(x) - h_t(x_0) \leq (x - x_0)\frac{x_0^{1/t-1}}{t}, \qquad x \geq x_0,$$

$$0 \leq h_t(x_0) - h_t(x) \leq (x_0 - x)\, x_0^{1/t-1}, \qquad x \leq x_0,$$

from which we deduce the following inequality: for all

$$(h_t(x) - h_t(x_0))^2 \leq (x - x_0)^2 \left(x_0^{1/t-1}\max\left\{1, \frac{1}{t}\right\}\right)^2.$$

Now returning to the estimation of the lead eigenvalue, by appealing to the above inequality and $(x + y)^2 \leq 2x^2 + 2y^2$ we get

$$
\mathbb{E}\left[\left(\left(\Psi_k^{\mathtt{br}}[g](t, r, \upsilon)\right)^{1/t} - \mathrm{e}^{\lambda_*}\right)^2\right]
$$

(C.7)

$$
\leq \frac{2}{\min\{1, t^2\}} \frac{\mathbb{E}[(\Psi_k^{\mathtt{br}}[g](t, r, \upsilon) - \psi_t[g](r, \upsilon))^2]}{\psi_t[g](r, \upsilon)^{2 - 2/t}} + 2\left((\psi_t[g](r, \upsilon))^{1/t} - \mathrm{e}^{\lambda_*}\right)^2.
$$

The proof of Theorem 5.1 now follows from the above inequality, (3.2), and the estimates in (i)–(iii) of Lemma C.1. ∎

Appendix D. NRW Monte Carlo convergence. We need the following intermediate result, before turning to the proof of Theorem 6.1.

Lemma D.1. *Under the conditions of Theorem* 6.1, *we have*

(D.1)
$$
\max(2\lambda_*, \lambda_* + \underline{\beta}) \leq \lambda_1 \leq \lambda_* + \overline{\beta}.
$$

Proof. We have, on the one hand, that

$$
\psi_t^1[g^2](r, \upsilon) := \mathbf{E}_{(r, \upsilon)}\left[\mathrm{e}^{\int_0^t 2\beta(R_s, \Upsilon_s)\mathrm{d}s} g(R_t, \Upsilon_t)^2 \mathbf{1}_{(t < \tau^D)}\right]
$$

(D.2)
$$
\leq \mathrm{e}^{\bar{\beta}t}\mathbf{E}_{(r, \upsilon)}\left[\mathrm{e}^{\int_0^t \beta(R_s, \Upsilon_s)\mathrm{d}s} g(R_t, \Upsilon_t)^2 \mathbf{1}_{(t < \tau^D)}\right] = \mathrm{e}^{\bar{\beta}t}\psi_t[g^2](r, \upsilon).
$$

On the other hand, by Jensen's inequality, we have that

(D.3)
$$
\psi_t^1[g^2](r, \upsilon) \geq \mathbf{E}_{(r, \upsilon)}\left[\mathrm{e}^{\int_0^t \beta(R_s, \Upsilon_s)\mathrm{d}s} g(R_t, \Upsilon_t)\mathbf{1}_{(t < \tau^D)}\right]^2 = \psi_t[g](r, \upsilon)^2
$$

and also that

(D.4)
$$
\psi_t^1[g^2](r, \upsilon) \geq \mathrm{e}^{\underline{\beta}t}\mathbf{E}_{(r, \upsilon)}\left[\mathrm{e}^{\int_0^t \beta(R_s, \Upsilon_s)\mathrm{d}s} g(R_t, \Upsilon_t)^2 \mathbf{1}_{(t < \tau^D)}\right]^2 = \mathrm{e}^{\underline{\beta}t}\psi_t[g^2](r, \upsilon).
$$

By taking logarithms, dividing by t, and then taking $t \to \infty$ using Theorem 3.1 in (D.2) and (D.3), (D.4), respectively, yields the desired bounds. ∎

Proof of Theorem 6.1. In a similar spirit to (C.1), we note that

$$
\mathbf{E}_{(r, \upsilon)}\left[\left(\Psi_k^{\mathtt{rw}}[g](t, r, \upsilon) - \psi_t[g](r, \upsilon)\right)^2\right]
$$

(D.5)
$$
= \frac{1}{k}\left(\mathbf{E}_{(r, \upsilon)}\left[\mathrm{e}^{\int_0^t 2\beta(R_s, \Upsilon_s)\mathrm{d}s} g(R_t, \Upsilon_t)^2 \mathbf{1}_{(t < \tau^D)}\right] - \psi_t[g](r, \upsilon)^2\right).
$$

Next we note that $\psi_t^1[g^2]$, which was previously defined in (D.2), behaves similarly to $\psi_t[g]$, albeit with the potential β replaced by 2β and g by g^2. Invoking Theorem 3.1, taking note of Remark 2.2, we thus conclude that there exists a λ_1 with accompanying eigenfunctions φ_1 and $\tilde{\varphi}_1$ in $L_\infty^+(D \times V)$ such that

(D.6)
$$
\sup_{g \in L_\infty^+(D \times V): \|g\|_\infty \leq 1} \left\|\mathrm{e}^{-\lambda_1 t}(\varphi_1)^{-1}\psi_t^1[g^2] - \langle \tilde{\varphi}_1, g^2\rangle\right\|_\infty = o(\mathrm{e}^{-\varepsilon t}) \text{ as } t \to +\infty.
$$

Taking advantage of (C.7), the desired conclusion follows with the help of (D.6) and Lemma D.1. ∎

Appendix E. h-NRW Monte Carlo convergence.

Proof of Theorem 8.1. In that case, we note from the proof of Theorem 6.1 that the crux of the argument there boils down to the estimate (C.7) together with the analogue of (D.5), which reads

$$
\mathbf{E}^h_{(r,v)}\left[\left(\Psi^{h\text{-}\mathtt{rw}}_k[g](r,v) - \psi_t[g](r,v)\right)^2\right]
$$

$$
= \frac{1}{k}\left(\mathbf{E}^h_{(r,v)}\left[\exp\left(2\int_0^t \frac{\mathrm{L}h(R_s,\Upsilon_s)}{h(R_s,\Upsilon_s)} + \beta(R_s,\Upsilon_s)\mathrm{d}s\right)\frac{g(R_t,\Upsilon_t)^2}{h(R_t,\Upsilon_t)^2}\mathbf{1}_{(t<\tau^D)}\right] - \psi_t[g](r,v)^2\right)
$$

$$
\leq \frac{1}{k}\left(\mathbf{E}^h_{(r,v)}\left[\exp\left(2\int_0^t \frac{\mathrm{L}h(R_s,\Upsilon_s)}{h(R_s,\Upsilon_s)} + \beta(R_s,\Upsilon_s)\mathrm{d}s\right)\mathbf{1}_{(t<\tau^D)}\right] - \psi_t[g](r,v)^2\right)
$$

(E.1)
$$
= \frac{1}{k}\left(\mathbf{E}_{(r,v)}\left[\exp\left(\int_0^t \frac{\mathrm{L}h(R_s,\Upsilon_s)}{h(R_s,\Upsilon_s)} + 2\beta(R_s,\Upsilon_s)\mathrm{d}s\right)\frac{h(R_t,\Upsilon_t)}{h(r,v)}\mathbf{1}_{(t<\tau^D)}\right] - \psi_t[g](r,v)^2\right),
$$

where we have used that $g \leq h$.

To complete the proof, appealing to (8.3), we can follow the reasoning in the proof of Theorem 6.1 and Lemma D.1. In particular, if we write

$$
\overline{\varsigma} := \sup_{r\in D, v\in V} \frac{(\mathrm{L}+\beta)h(r,v)}{h(r,v)} \leq \sup_{r\in D, v\in V} \frac{\mathrm{L}h(r,v)}{h(r,v)} + \overline{\beta} < \infty
$$

and

$$
\underline{\varsigma} := \inf_{r\in D, v\in V} \frac{(\mathrm{L}+\beta)h(r,v)}{h(r,v)},
$$

then there exists a λ_2 satisfying $\lambda_* + \underline{\varsigma} \leq \lambda_2 \leq \lambda_* + \overline{\varsigma}$. such that the desired result holds. ■

Proof of Lemma 8.2. We can take inspiration from the proof of Lemma 6.2, which is given immediately preceding the statement of the lemma. In particular considering (6.2), when we are under the assumptions of Theorem 8.1, appealing to (8.2), for $f \in L^+_\infty(D \times V)$ the cost function (6.1) satisfies

(E.2) $\quad \mathbf{E}^h_{(r,v)}[C_t[f]] = \mathbf{E}^h_{(r,v)}\left[\int_0^t \mathbf{1}_{(s<\tau^D)}\alpha_h\pi_h[f](R_s,\Upsilon_s)\mathrm{d}s\right]$

$$
= \mathbf{E}^h_{(r,v)}\left[\int_0^t \mathbf{1}_{(s<\tau^D)}\alpha(R_s,\Upsilon_s)\int_V f(R_s,v')\frac{h(R_s,v')}{h(R_s,\Upsilon_s)}\pi(R_s,\Upsilon_s,v')\mathrm{d}v'\mathrm{d}s\right]
$$

$$
= \frac{1}{h(r,v)}\mathbf{E}_{(r,v)}\left[\int_0^t \mathbf{1}_{(s<\tau^D)}\mathrm{e}^{-\int_0^s \frac{\mathrm{L}h}{h}(R_u,\Upsilon_u)\mathrm{d}u}\alpha(R_s,\Upsilon_s)\right.
$$

$$
\left.\int_V f(R_s,\Upsilon_s)h(R_s,v')\pi(R_s,\Upsilon_s,v')\mathrm{d}v'\mathrm{d}s\right]
$$

(E.3)
$$
\leq \frac{\|\alpha h f\|_\infty}{h(r,v)}\int_0^t \mathbf{E}_{(r,v)}\left[\mathrm{e}^{\int_0^s \beta(R_u,\Upsilon_u)\mathrm{d}u}\mathbf{1}_{(s<\tau^D)}\mathrm{e}^{-\int_0^s \frac{(\mathrm{L}+\beta)h}{h}(R_u,\Upsilon_u)\mathrm{d}u}\right]\mathrm{d}s,
$$

where we have used (7.16). Recalling the definition in (8.5), as well as the conclusion of Theorem 3.1, it is straightforward to estimate from (E.3) that

$$\limsup_{t\to\infty} e^{-(\lambda_*-\varsigma)t}\mathbf{E}^h_{(r,v)}[C_t[f]] < \infty.$$

In the event that $\lambda_* = \varsigma$, we need to be a little more careful and note that the integrand is asymptotically a constant, and, in that case, the expected cost grows no faster than linearly.

In the second case, where the conditions on the domain are satisfied, then the argument follows by careful consideration of (E.2). We first define $U^f_n := \{(r,v)|\alpha_h\pi_h[f] \in [n, n+1)\}$. Note that as a consequence of the lower bound on f, scatter events on this set happen with rate at least nc_0^{-1}. Moreover, given there is a scatter event in U^f_n, the scatter event will result in scatter *to* Ω_0 with probability at least p_0. For a fixed time s, we write $\tau_s := \sup\{u \le s|(R_u, \Upsilon_u) \ne (R_s - (s-u)\Upsilon_s, \Upsilon_s)$, the time of the most recent scatter event. Then let

$$A^n_t := \{(R_s, \Upsilon_s) \in \Omega_0 \text{ for all } s \in [\tau_t, t], \text{ and } (R_{\tau_t-}, \Upsilon_{\tau_t-}) \in U^f_n\},$$

that is, the event in which, at time t, the process is in Ω_0, and has been in Ω_0 since the most recent scatter event, which happened in the set U^f_n. Note that the sets A^n_t are disjoint for fixed t.

As a consequence of the assumption on α_h, we can find $\epsilon = \epsilon(\delta) > 0$ such that, conditional on starting at any $(r,v) \in \Omega_0$, then the expected time to the next jump is at least ϵ. Then we have the (approximate!) inequality

$$\mathbb{E}^h_{(r,v)}\left[\int_0^{t\wedge\tau^D} \mathbf{1}_{A^n_s}\,\mathrm{d}s\right] \ge \mathbb{E}^h_{(r,v)}\left[\int_0^{t\wedge\tau_D} \mathbf{1}_{(R_s,\Upsilon_s)\in U^f_n}\,\mathrm{d}s\right](1 + \epsilon c_0^{-1}p_0 n),$$

where the identity follows from the fact that the process jumps out of the set U^f_n into Ω_0 at rate at least $c_0^{-1}p_0 n$, and once it is there, spends on average at least ϵ time units.

Since we must have

$$\sum_{n\ge 0} \mathbb{E}^h_{(r,v)}\left[\int_0^{t\wedge\tau^D} \mathbf{1}_{A^n_s}\,\mathrm{d}s\right] \le t$$

we conclude that

$$\sum_{n\ge 0} n\cdot\mathbb{E}^h_{(r,v)}\left[\int_0^{t\wedge\tau^D} \mathbf{1}_{(R_s,\Upsilon_s)\in U^f_n}\,\mathrm{d}s\right] \le Ct$$

for some constant C. Substituting this estimate into (E.2) gives the desired conclusion. ∎

Appendix F. The one-dimensional slab reactor.

Proof of Proposition 9.1. Denote by Λ the set of eigenvalues for which there exist $\tilde\varphi \in L^+_\infty(D \times V)$ and $\varphi \in L^+_\infty(D \times V)$ such that (9.3) holds. According to Theorem 3.1, $\lambda_* \in \Lambda$ and $\lambda_* = \sup \Lambda$.

Suppose that $\lambda \in \Lambda$, and let χ be the associated eigenfunction: $A\chi = \lambda\chi$. Let us write $f_+(r) = \chi(r, v_0)$ and $f_-(r) = \chi(r, -v_0)$. Using (9.2), we can build candidates for χ in $L_2^+(D \times V)$ from the pointwise relations

$$(\text{F.1}) \quad \frac{d}{dx}\begin{pmatrix} f_+ \\ f_- \end{pmatrix} = M_\lambda \begin{pmatrix} f_+ \\ f_- \end{pmatrix}, \quad \text{where } M_\lambda = \frac{1}{v_0}\begin{pmatrix} \lambda - \sigma_\mathtt{f} + \sigma_\mathtt{s} & -\sigma_\mathtt{s} \\ \sigma_\mathtt{s} & -(\lambda - \sigma_\mathtt{f} + \sigma_\mathtt{s}) \end{pmatrix},$$

together with the boundary conditions

$$(\text{F.2}) \qquad\qquad\qquad f_+(L) = f_-(-L) = 0.$$

To solve (F.1) and (F.2), let us first suppose that

$$(\text{F.3}) \qquad\qquad (\lambda - \sigma_\mathtt{f} + \sigma_\mathtt{s})^2 \neq \sigma_\mathtt{s}^2 \quad \Longleftrightarrow \quad \lambda \neq \sigma_\mathtt{f} \text{ or } \lambda \neq \sigma_\mathtt{f} - 2\sigma_\mathtt{s}.$$

In this case, M_λ has two distinct eigenvalues, $\pm\alpha_\lambda$, where

$$(\text{F.4}) \qquad\qquad \alpha_\lambda = \sqrt{-\text{Det}(M_\lambda)} = \frac{1}{v_0}\sqrt{(\lambda - \sigma_\mathtt{f} + \sigma_\mathtt{s})^2 - \sigma_\mathtt{s}^2} \in \mathbb{R}_+ \cup \mathtt{i}\mathbb{R}_+.$$

Then by elementary results for ODEs, for fixed λ, the solutions to (F.1), (F.2) take the form

$$(\text{F.5}) \qquad f_+(r) = C(\mathrm{e}^{-\alpha_\lambda r} - \mathrm{e}^{\alpha_\lambda r - 2\alpha_\lambda L}), \qquad f_-(r) = C'(\mathrm{e}^{\alpha_\lambda r} - \mathrm{e}^{-\alpha_\lambda r - 2\alpha_\lambda L}),$$

where $C, C' \in \mathbb{R} \setminus \{0\}$. Differentiating (F.5), we find

$$\frac{d}{dx}\begin{pmatrix} f_+ \\ f_- \end{pmatrix} = \begin{pmatrix} -\alpha_\lambda \coth(2\alpha_\lambda L) & -\frac{C}{C'}\frac{\alpha_\lambda}{\sinh(2\alpha_\lambda L)} \\ \frac{C'}{C}\frac{\alpha_\lambda}{\sinh(2\alpha_\lambda L)} & \alpha_\lambda \coth(2\alpha_\lambda L) \end{pmatrix}\begin{pmatrix} f_+ \\ f_- \end{pmatrix}.$$

Comparing this with (F.1), we get

$$(\text{F.6}) \qquad \frac{C}{C'} = \frac{C'}{C}, \quad \frac{C'}{C}\frac{\alpha_\lambda}{\sinh(2\alpha_\lambda L)} = \frac{\sigma_\mathtt{s}}{v_0}, \quad \text{and} \quad \alpha_\lambda \coth(2\alpha_\lambda L) = -\frac{\lambda - \sigma_\mathtt{f} + \sigma_\mathtt{s}}{v_0}.$$

This first identity yields $C/C' = 1$ or -1. This then implies

$$(\text{F.7}) \qquad\qquad\qquad \left|\frac{2\alpha_\lambda}{\sinh(2\alpha_\lambda L)}\right| = \frac{2\sigma_\mathtt{s}}{v_0}.$$

Recall that $\alpha_\lambda \in \mathbb{R}_+ \cup \mathtt{i}\mathbb{R}_+$. If $\alpha_\lambda \in \mathbb{R}_+$, then (F.7) has a solution if and only if $\theta = v_0/(2\sigma_\mathtt{s}) \geq 1$. Moreover, if $\theta > 1$, (F.7) has a unique solution in $[0, \infty)$. Denote by x_* this unique solution. It follows from (F.7) that $\alpha_\lambda = \frac{x_*}{2L}$. Since $x/\sinh(x) \geq 0$ for $x \in \mathbb{R}$, (F.6) implies that

$$C = C' \quad \text{and} \quad \lambda - \sigma_\mathtt{f} + \sigma_\mathtt{s} < 0.$$

We then deduce from (F.4) that in the case where $\theta > 1$, we have

$$\alpha_\lambda = \frac{x_*}{2L} \quad \text{and} \quad \lambda = \sigma_\mathtt{f} - \sigma_\mathtt{s} - \sqrt{\sigma_\mathtt{s}^2 + \left(\frac{v_0 x_*}{2L}\right)^2}.$$

Plugging this into (F.5) and renormalizing f_+, f_- by taking $C^{-1} = f_+(0)$, we obtain the formulas (9.5) and (9.6) in the case $\theta > 1$. Next, suppose that $\alpha_\lambda \in i\mathbb{R}_+$. In that case, (F.7) has a solution if and only if $\theta = v_0/(2\sigma_{\mathtt{s}}) \leq 1$. Moreover, if $\theta < 1$, (F.7) has a finite number of solutions in $i\mathbb{R}_+$, which we denote by $\mathtt{i}x_k$, $1 \leq k \leq n$, with $0 < x_1 < x_2 < \cdots < x_n$. By a similar argument as before, for each x_k, $1 \leq k \leq n$, we get

$$\alpha_\lambda = \frac{x_k \mathtt{i}}{2L}, \quad \lambda = \sigma_{\mathtt{f}} - \sigma_{\mathtt{s}} - \mathrm{sgn}(\cot(x_k))\sqrt{\sigma_{\mathtt{s}}^2 - \left(\frac{v_0 x_k}{2L}\right)^2}, \quad \text{and} \quad \frac{C}{C'} = \mathrm{sgn}(\cot(x_k)).$$

Moreover, we have $x_1 \in (0, \pi)$. Therefore, in the case $\theta < 1$, $\lambda_* = \max \Lambda$ is given by (9.10) and we deduce the forms of the corresponding $\varphi, \tilde{\varphi}$ from (F.5).

Finally, we consider the case where $(\lambda - \sigma_{\mathtt{f}} + \sigma_{\mathtt{s}})^2 = \sigma_{\mathtt{s}}^2$. An elementary computation shows that (F.1) and (F.2) have a solution in this case if and only if $\theta = 1$. In that case, the solution is given by

$$\lambda = \sigma_{\mathtt{f}} - 2\sigma_{\mathtt{s}}, \qquad f_+(r) = 1 - \frac{r}{L}, \quad \text{and} \quad f_-(r) = 1 + \frac{r}{L}.$$

Together with our earlier remarks regarding the relationship between φ and $\tilde{\varphi}$, this concludes the proof. ■

REFERENCES

[1] L. ANGELI, S. GROSSKINSKY, AND A. M. JOHANSEN, *Limit Theorems for Cloning Algorithms*, preprint, http://arxiv.org/abs/1902.00509, 2019 (accessed 2019-11-22).

[2] G. I. BELL, *On the stochastic theory of neutron transport*, Nucl. Sci. Eng., 21 (1965), pp. 390–401.

[3] G. I. BELL AND S. GLASSTONE, *Nuclear Reactor Theory*, Reinhold, 1970.

[4] M. BENAIM, B. CLOEZ, AND F. PANLOUP, *Stochastic approximation of quasi-stationary distributions on compact spaces and applications*, Ann. Appl. Probab., 28 (2018), pp. 2370–2416.

[5] J. BERTOIN, *Random Fragmentation and Coagulation Processes*, Cambridge Stud. Adv. Math. 102, Cambridge University Press, Cambridge, 2006.

[6] F. B. BROWN, *Monte Carlo Techniques for Nuclear Systems - Theory Lectures*, Tech. Report LA-UR-16-29043, Los Alamos National Lab (LANL), Los Alamos, NM, 2016, https://doi.org/10.2172/1334102.

[7] E. BRUN, F. DAMIAN, C. M. DIOP, E. DUMONTEIL, F. X. HUGOT, C. JOUANNE, Y. K. LEE, F. MALVAGI, A. MAZZOLO, O. PETIT, J. C. TRAMA, T. VISONNEAU, AND A. ZOIA, TRIPOLI-4, CEA, *EDF and AREVA reference Monte Carlo code*, Ann. Nucl. Energy, 82 (2015), pp. 151–160, https://doi.org/10.1016/j.anucene.2014.07. 053.

[8] K. BURDZY, R. HOLYST, AND P. MARCH, *A Fleming-Viot particle representation of the Dirichlet Laplacian*, Commun. Math. Phys., 214 (2000), pp. 679–703, https://doi.org/10.1007/s002200000294.

[9] N. CHAMPAGNAT AND D. VILLEMONAIS, *Convergence of the Fleming-Viot Process toward the Minimal Quasi-stationary Distribution*, preprint, http://arxiv.org/abs/1810.06849, 2018.

[10] Z.-Q. CHEN AND R. SONG, *Conditional gauge theorem for non-local Feynman-Kac transforms*, Probab. Theory Related Fields, 125 (2003), pp. 45–72, https://doi.org/10.1007/s004400200219.

[11] S. Christoforou and J. E. Hoogenboom, *A zero-variance-based scheme for Monte Carlo criticality calculations*, Nucl. Sci. Eng., 167 (2011), pp. 91–104, https://doi.org/10.13182/NSE09-107.

[12] K. L. Chung and J. B. Walsh, *Markov Processes, Brownian Motion, and Time Symmetry*, 2nd ed., Grundlehren Math. Wiss. 249, Springer, New York, 2005, https://doi.org/10.1007/0-387-28696-9.

[13] A. M. G. Cox, S. C. Harris, E. L. Horton, and A. E. Kyprianou, *Multi-species neutron transport equation*, J. Stat. Phys., 176 (2019), pp. 425–455, https://doi.org/10.1007/s10955-019-02307-2.

[14] R. Dautray, M. Cessenat, G. Ledanois, P.-L. Lions, E. Pardoux, and R. Sentis, *Méthodes probabilistes pour les équations de la physique*, Collection du Commissariat a l'énergie atomique, Eyrolles, Paris, 1989.

[15] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 6, Evolution Problems* II, Springer-Verlag, Berlin, 1993.

[16] M. H. A. Davis, *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, J. Roy. Statist. Soc. Ser. B, 46 (1984), pp. 353–376, https://doi.org/10.1111/j.2517-6161.1984.tb01308.x.

[17] B. Davison and J. B. Sykes, *Neutron Transport Theory*, Clarendon Press, Oxford, 1957.

[18] P. Del Moral, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications, Probab. Appl.,* Springer-Verlag, New York, 2004.

[19] R. Douc, O. Cappé, and E. Moulines, *Comparison of resampling schemes for particle filtering*, in Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis (ISPA 2005), IEEE, 2005, pp. 64–69, https://doi.org/10.1109/ISPA.2005.195385.

[20] A. Doucet, N. d. Freitas, and N. Gordon, eds., *Sequential Monte Carlo Methods in Practice, Information Science and Statistics*, Springer-Verlag, New York, 2001, https://doi.org/10.1007/978-1-4757-3437-9.

[21] E. B. Dynkin, *Diffusions, Superdiffusions and Partial Differential Equations*, Amer. Math. Soc. Colloq. Publ. 50, American Mathematical Society, Providence, RI, 2002.

[22] A. M. Etheridge, *An Introduction to Superprocesses*, Univ. Lecture Ser. 20, American Mathematical Society, Providence, RI, 2000.

[23] M. Faucher, D. Mancusi, and A. Zoia, *Variance-reduction methods for Monte Carlo kinetic simulations*, in International Conference on Mathematics and Computational Methods (Portland, OR, 2019), American Nuclear Society, 2019.

[24] I. G. Graham, M. J. Parkinson, and R. Scheichl *Modern Monte Carlo variants for uncertainty quantification in neutron transport*, in Festschrift for the 80th Birthday of Ian Sloan, J. Dick, F. Y. Kuo, and H. Wozniakowski, eds., Springer, Cham, 2018, pp. 455–481.

[25] P. Groisman and M. Jonckheere, *Simulation of Quasi-stationary Distributions on Countable Spaces*, preprint, http://arxiv.org/abs/1206.6712, 2012.

[26] S. C. Harris, E. Horton, and A. E. Kyprianou, *Stochastic methods for the neutron transport equation II: Almost sure growth*, Ann. Appl. Probab., 30 (2020), pp. 2815–2845, https://doi.org/10.1214/20-AAP1574.

[27] S. C. Harris and M. I. Roberts, *The many-to-few lemma and multiple spines*, Ann. Inst. Henri Poincaré Probab. Stat., 53 (2017), pp. 226–242, https://doi.org/10.1214/15-AIHP714.

[28] J. Heng, A. N. Bishop, G. Deligiannidis, and A. Doucet, *Controlled sequential Monte Carlo*, Ann. Statist., 48 (2020), pp. 2904–2929, https://doi.org/10.1214/19-AOS1914.

[29] E. Horton, A. E. Kyprianou, and D. Villemonais, *Stochastic methods for the neutron transport equation I: Linear semigroup asymptotics*, Ann. Appl. Probab., 30 (2020), pp. 2573–2612, https://doi.org/10.1214/20-AAP1567.

[30] R. L. Jack, *Ergodicity and large deviations in physical systems with stochastic dynamics*, Eur. Phys. J. B, 93, (2020), 74.

[31] A. E. Kyprianou and S. Palau, *Extinction properties of multi-type continuous-state branching processes*, Stochastic Process. Appl., 128 (2018), pp. 3466–3489, https://doi.org/10.1016/j.spa.2017.11.006.

[32] B. Lapeyre, E. Pardoux, and R. Sentis, *Introduction to Monte-Carlo Methods for Transport and Diffusion Equations*, Oxf. Texts Appl. Eng. Math. 6, Oxford University Press, Oxford, 2003.

[33] J.-F. Le Gall, *Random trees and applications*, Probab. Surv., 2 (2005), pp. 245–311, https://doi.org/10.1214/154957805100000140.

[34] J. Leppänen, M. Pusa, T. Viitanen, V. Valtavirta, and T. Kaltiaisenaho, *The Serpent Monte Carlo code: Status, development and applications in* 2013, Ann. Nucl. Energy, 82 (2015), pp. 142–150, https://doi.org/10.1016/j.anucene.2014.08. 024.

[35] E. E. Lewis and J. W. F. Miller, *Computational Methods of Neutron Transport*, Wiley & Sons, New York, 1984.

[36] Z. Li, *Measure-Valued Branching Markov Processes*, Probab. Appl. (N.Y.), Springer, Heidelberg, 2011, https://doi.org/10.1007/978-3-642-15004-3.

[37] I. Lux and L. Koblinger, Monte Carlo Particle Transport Methods: Neutron and Photon Calculations, CRC Press, 1991, https://doi.org/10.1201/9781351074834.

[38] C. Mailler and D. Villemonais, *Stochastic approximation on noncompact measure spaces and application to measure-valued Pólya processes*, Ann. Appl. Probab., 30 (2020), pp. 2393–2438, https://doi.org/10.1214/20-AAP1561.

[39] S. Maire and D. Talay, *On a Monte Carlo method for neutron transport criticality computations*, IMA J. Numer. Anal., 26 (2006), pp. 657–685.

[40] D. Mancusi and A. Zoia, *Towards zero-variance schemes for kinetic Monte-Carlo simulations*, EPJ Web. Conf., 247 (2021), 04010, https://doi.org/10.1051/epjconf/202124704010.

[41] M. Mokhtar-Kharroubi, *Mathematical Topics in Neutron Transport Theory*, New Aspects, Ser. Adv. Math. Appl. Sci. 46, World Scientific, River Edge, NJ, 1997.

[42] I. Pázsit and L. Pál, *Neutron Fluctuations: A Treatise on the Physics of Branching Processes*, Elsevier, 2008.

[43] A. Pazy and P. H. Rabinowitz, *On a branching process in neutron transport theory*, Arch. Rational Mech. Anal., 51 (1973), pp. 153–164.

[44] M. Pollock, P. Fearnhead, A. M. Johansen, and G. O. Roberts, *Quasi-stationary Monte Carlo and the ScaLE Algorithm*, preprint, http://arxiv.org/abs/1609.03436, 2020.

[45] S. Richards, G. Dobson, T. Fry, D. Hanlon, D. Long, R. Perry, P. Smith, F. Tantillo, and T. Ware. *Recent developments to the MONK Monte Carlo code for criticality safety and reactor physics analyses*, in ICNC, 2019 - 11th International conference on Nuclear Criticality Safety, 2019.

[46] P. Vanetti, A. Bouchard-Côté, G. Deligiannidis, and A. Doucet, *Piecewise-Deterministic Markov Chain Monte Carlo*, preprint, http://arxiv.org/abs/1707.05296, 2018.

[47] A. Q. Wang, G. O. Roberts, and D. Steinsaltz, *An approximation scheme for quasi-stationary distributions of killed diffusions*, Stochastic Process. Appl., 130 (2020), pp. 3193–3219, https://doi.org/10.1016/j.spa.2019.09.010.

[48] C. J. Werner, J. S. Bull, C. J. Solomon, F. B. Brown, G. W. McKinney, M. E. Rising, D. A. Dixon, R. L. Martz, H. G. Hughes, L. J. Cox, A. J. Zukaitis, J. C. Armstrong, R. A. Forster, and L. Casswell, *MCNP Version 6.2 Release Notes*, Tech. Report LA-UR-18-20808, Los Alamos National Lab (LANL), Los Alamos, NM, 2018, https://doi.org/10.2172/1419730.

[49] N. Whiteley and A. Lee, *Twisted particle filters*, Ann. Statist., 42 (2014), pp. 115–141, https://doi.org/10.1214/13-AOS1167

[50] A. Zoia, E. Brun, and F. Malvagi, *Alpha eigenvalue calculations with Tripoli-4*, Ann. Nucl. Energy, 63 (2014), pp. 276–284, https://doi.org/10.1016/j.anucene.2013.07.018.