

Stochastic Analysis of the Neutron Transport Equation

submitted by

Emma Louise Horton

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

November 2019

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with the author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that they must not copy it or use material from it except as permitted by law or with the consent of the author.

This thesis may be made available for consultation
within the University Library and may be
photocopied or lent to other libraries for the purposes
of consultation with effect from (date)

Signed on behalf of the Faculty of Science

Declaration of any previous Submission of the Work

The material presented here for examination for the award of a higher degree by research has not been incorporated into a submission for another degree.

.....

Emma Louise Horton

Declaration of Authorship

I am the author of this thesis, and the work described therein was carried out by myself personally, in collaboration with my supervisors Alexander M. G. Cox, Simon C. Harris and Andreas E. Kyprianou, and external collaborator Denis Villemonais.

.....

Emma Louise Horton

Summary

This thesis concerns the study of the neutron transport equation, which describes how neutrons move through a fissile medium, such as a nuclear reactor. By studying various stochastic processes that model the behaviour of these neutrons, we address some of the criticality problems associated with such systems.

We first build a class of branching processes, whose behaviour mimics that of neutrons undergoing fission in a nuclear reactor. We then construct a class of weighted random walks, which are in some sense equivalent to the branching processes via a many-to-one formula. Analysis of these two processes allows us to characterise the long-term behaviour of the underlying nuclear fission processes.

One of the parameters associated with characterising this asymptotic behaviour quantifies the average growth of particle numbers. A large part of this thesis is dedicated to what is known as the supercritical phase, which is where the average number of particles grows exponentially. In this regime, we consider both a spine and skeletal decomposition of the branching process in order to further describe the growth of the system through a law of large numbers result.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Andreas Kyprianou for his support, patience and enthusiasm during my time at Bath as both a PhD student and an undergraduate. I have thoroughly enjoyed working together these past six years and I am truly grateful for all the wonderful experiences and opportunities you have given me. I would also like to thank my supervisors Alex Cox and Simon Harris for their guidance, encouragement and numerous enjoyable meetings. It has been a pleasure and a privilege to work with three such inspiring researchers, who always found time for me during their increasingly busy schedules.

Besides my supervisors, I would like to thank my collaborator Denis Villemonais, from whom I have learnt a lot. Thank you for always finding time to host me in Nancy, for the many fruitful discussions and, of course, for giving me my first postdoc position. A huge thank you also goes to Paul, Geoff and the rest of the staff at Wood for always making us welcome at their offices in Dorchester. I enjoyed every minute of these visits and always returned to Bath with several new ideas after our many stimulating discussions. I would also like to say a massive thank you to Jess and Susie; their kindness, support and the well-stocked selection of gluten free cakes and biscuits made my SAMBa experience truly amazing.

A very special thank you goes to Katherine, Kayley and Sam for their friendship and the many wonderful adventures that kept me going through the highs and lows of my PhD. To Jess, Rosie, Georgie, Rosanna, Mel, Mary and Ali, thank you for always being there to dance the stress away and for generally making my time in Bath unforgettable. I also wish to thank Dorka, Cécile, Cheryl, Isaac, Minmin, Cameron, Nadeen and Tsoogii for their advice, friendship and for always being there for coffee when I needed it.

Finally, I would like to thank my parents and my sisters, Sarah and Lucy, for their unconditional love and endless support. I honestly could not have done this without you.

Contents

1	Introduction	1
1.1	The physical process	1
1.1.1	Transport	1
1.1.2	Scattering	2
1.1.3	Fission	2
1.2	Neutron Transport Equation	3
1.2.1	Forwards equation	3
1.2.2	Backwards equation	4
1.3	Criticality problems	5
1.4	Monte Carlo methods	7
1.5	Notation	10
1.6	Outline of thesis	11
2	Multi-species neutron transport equation	18
2.1	Introduction	19
2.2	Neutron Transport Equation	20
2.3	Organisation of the paper	25
2.4	Historical remarks on the mathematical treatment of the NTE	26
2.5	Multi-species (Backwards) Neutron Transport Equation	27
2.6	Multi-species neutron branching process	33
2.7	Multi-species neutron random walk and the Many-to-one Lemma	38
2.8	Consolidating the ACP with the expectation semigroup	41
2.9	Asymptotic behaviour of the MNTE: Proof of Theorem 3	44
3	Linear semigroup asymptotics	55
3.1	Introduction	55
3.2	The physical process and the mild NTE	57
3.3	Perron-Frobenius asymptotics	61
3.4	Neutron random walk and many-to-one methodology	64
3.5	The ground state martingale	66
3.6	Neutron random walk and spine decomposition	67
3.7	Proof of Theorem 7	69

3.8	Proof of Theorem 9	85
3.9	Proof of Lemma 5	90
3.10	Proof of Theorem 8	92
3.11	Proof of Corollary 2	96
4	Skeleton decompositions and a strong law of large numbers	104
4.1	Introduction	105
4.1.1	Rigorous interpretation of the NTE	106
4.1.2	Neutron Branching Process	107
4.1.3	Lead order asymptotics of the expectation semigroup	110
4.1.4	Strong law of large numbers at supercriticality	111
4.2	Skeletal decomposition	112
4.2.1	The general branching Markov setup	113
4.2.2	\downarrow -trees	116
4.2.3	\uparrow -trees	119
4.2.4	Combining \uparrow -trees and \downarrow -trees into the skeletal decomposition	124
4.2.5	Remarks on the skeletal decomposition for the NBP	125
4.2.6	Remarks on BBM	126
4.3	SLLN on the skeleton	127
4.4	Proof of Theorem 13	135
4.5	Concluding remarks	136
5	Time-independent criticality problems	144
5.1	Introduction	145
5.2	Formulations of the NTE and associated eigenfunctions	148
5.2.1	Abstract Cauchy Problem (ACP)	148
5.2.2	Neutron branching process (NBP) and the mild NTE	149
5.2.3	Neutron random walk (NRW)	151
5.2.4	Neutron generational process (NGP)	152
5.3	Probabilistic solution to (5.6)	153
5.4	Classical existence of solution to (5.6)	156
5.5	Proof of Theorem 19	159
5.6	Concluding remarks	162
5.6.1	λ -, k - and c -eigenvalue problems	162
5.6.2	Martingale convergence and strong law of large numbers	163
5.6.3	Monte-Carlo considerations	163
6	Conclusions	168

Chapter 1

Introduction

The Neutron Transport Equation (NTE) is used to model the movement of neutrons in fissile environments. This thesis concerns the development of probabilistic methods for studying the NTE in order to characterise the growth of such systems. We begin this chapter by describing some of the physics that governs the movement of neutrons in these environments.

1.1 The physical process

The neutron transport equation (NTE) describes the flux of neutrons in an inhomogeneous fissile medium, such as a nuclear reactor. It can be described as a function of time, t , Euclidian location, $r \in \mathbb{R}^3$, direction of travel, $\Omega \in \mathbb{S}_2$, and neutron energy, $E \in (0, \infty)$. However, it is usual to assume that energy is a function of velocity ($E = m|v|^2/2$), thereby reducing the number of variables by one. This allows us to describe the dependency of flux more simply in terms of time and the *configuration variables* $(r, v) \in D \times V$ where $D \subseteq \mathbb{R}^3$ is a smooth, open, connected and bounded domain¹ such that ∂D has zero Lebesgue measure and V is the set of admissible velocities, given by $V = \{v \in \mathbb{R}^3 : v_{\min} < |v| < v_{\max}\}$, where $0 < v_{\min} < v_{\max} < \infty$.

In order to describe the evolution of the flux, we describe the main processes that dictate how neutrons move in their domain.

1.1.1 Transport

The system is started from an initial source term, denoted $Q(r, v, t)$, with $t = 0$, $r \in D$ and $v \in V$. After this, a neutron with initial configuration (r, v) will move along the trajectory $(r + vs, v)$, $s \geq 0$, until it either hits the boundary of the reactor, at which point it is absorbed and no longer tracked, or either a scattering or fission event occurs, which we describe below. We note that in some radiation transport systems, the source term continues to contribute particles to the system throughout the whole process, however in most fissile media, the source term is neglected after a fixed period of time since the subsequent fission overwhelms this term.

¹We will later impose further assumptions on the domain. We note that in the literature, infinite domains have also been considered.

1.1.2 Scattering

A scattering event occurs when a neutron collides with a nucleus in the reactor and then “bounces” off, continuing to move in a straight line but with a new velocity. There are two main types of scattering, which we briefly describe.

Elastic scattering

This is the most common type of scattering. It occurs when the total energy involved in the scattering event is conserved, i.e. the nucleus involved in the collision gains precisely the amount of energy lost by the colliding neutron. In fact, elastic scattering can be further subdivided into compound and potential elastic scattering, depending on whether the incident neutron is absorbed by the nucleus or just comes into close proximity with it during the scattering event.

Inelastic scattering

Inelastic scattering only occurs when the energy of the colliding neutron is sufficiently high. In this case, the nucleus initially absorbs the energy lost by the colliding nucleus but it is left in an excited state. In order to return to stability, it releases energy in the form of gamma rays, so that the total energy of the neutron and nucleus after the collision is less than the energy beforehand.

In this thesis, we will not distinguish between the different types of scattering. However, the models we consider are general enough to allow a distinction to be made if necessary.

1.1.3 Fission

Roughly speaking, nuclear fission occurs when a neutron collides with a nucleus, which then causes the nucleus to release several more neutrons. Initially, the incident neutron is absorbed by the nucleus, which results in the nucleus becoming ‘excited’ or unstable. In order to return to its ground state, the nucleus splits open and releases surplus neutrons. However, not all neutrons are released from the nucleus at the same time, which leads us to the following two categories of fission neutrons.

Prompt neutrons

Neutrons that are released immediately (on a time scale of about 10^{-14} seconds) from a fission event are called prompt neutrons. Most of the neutron yield (about 99%) from fission events is made up of prompt neutrons, with the average yield of such neutrons being between two and three per fission event.

Delayed neutrons

When the excited nucleus splits open and releases prompt neutrons, the remaining fragments from the nucleus may still be in an excited state. These fragments are called *delayed neutron*

precursors and after some time, they release further neutrons, called delayed neutrons, in order to become stable again. Although delayed neutrons constitute a much lower proportion of the total yield from a fission event, they are very important in the design and safety of nuclear reactors. The time taken to release these neutrons can be anything from a few seconds to a minute, which is enough time to allow procedures to be put in place to absorb surplus neutrons in the case where too many are being produced, i.e. when the reactor is in a supercritical state.

It is possible that when a neutron is absorbed by a nucleus in a fission event, the resulting nucleus remains in a stable state. In this case, the nucleus will not release any neutrons, and we refer to this possibility as neutron *capture* or *absorption*. In the literature, capture and fission are often considered as two separate events; however it is not uncommon to consider absorption to be a fission event with zero offspring.

Although we have only described the behaviour of neutrons, there are several other types of particles involved in fission processes, such as gamma rays and alpha particles. Models that include these other various types of particles shall be discussed in Chapter 2.

1.2 Neutron Transport Equation

Based on the above dynamics, we are now in a position to write down the neutron transport equation. For now, we will only consider prompt neutrons, and as previously mentioned, we do not make a distinction between the different types of scattering.

1.2.1 Forwards equation

For $t \geq 0$, we denote the neutron flux at time t by $\Psi_t : D \times V \rightarrow [0, \infty)$. By considering the dynamics described above and balancing neutron production and loss, one can show, at least heuristically, that it solves the integro-differential equation, also known as the *forward neutron transport equation*²,

$$\begin{aligned} \frac{\partial}{\partial t} \Psi_t(r, v) &= -v \cdot \nabla \Psi_t(r, v) - \sigma(r, v) \Psi_t(r, v) \\ &+ \int_V \Psi_t(r, v') \sigma_s(r, v') \pi_s(r, v', v) dv' + \int_V \Psi_t(r, v') \sigma_f(r, v') \pi_f(r, v', v) dv', \end{aligned} \quad (1.1)$$

²Here and everywhere else in this thesis, ∇ is the gradient operator with respect to the variable $r \in \mathbb{R}^3$.

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

- $\sigma_{\mathbf{s}}(r, v')$: the rate at which scattering occurs from incoming velocity v' ,
- $\sigma_{\mathbf{f}}(r, v')$: the rate at which fission occurs from incoming velocity v' ,
- $\sigma(r, v)$: the sum of the rates $\sigma_{\mathbf{f}} + \sigma_{\mathbf{s}}$, also known as the *total* cross section
- $\pi_{\mathbf{s}}(r, v', v)dv'$: the scattering yield at velocity v from incoming velocity v' ,
satisfying $\int_V \pi_{\mathbf{s}}(r, v, v')dv' = 1$,
- $\pi_{\mathbf{f}}(r, v', v)dv'$: the neutron yield at velocity v from fission with incoming velocity v' ,
satisfying $\int_V \pi_{\mathbf{f}}(r, v, v')dv' < \infty$.

It is also natural to consider the following boundary conditions.

$$\begin{cases} \Psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \Psi_t(r, v) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } v \cdot \mathbf{n}_r < 0, \end{cases} \quad (1.2)$$

where \mathbf{n}_r is the outward facing normal of D at $r \in \partial D$ and $g : D \times V \rightarrow [0, \infty)$ is a bounded, measurable function. Heuristically, the forwards equation and corresponding boundary conditions describe possible previous configurations of neutrons in the reactor in order to achieve the current configuration.

The forwards equation is also sometimes written including the source term Q , which continually produces neutrons while the reactor is running. In the presence of a source term, the initial data is taken to be $g = Q$. However, as we will only be considering fissile systems, we have set $Q = 0$, as the resulting fission will overwhelm the radioactive source term.

1.2.2 Backwards equation

Although the forwards equation is the most common form of the NTE in the physics and engineering literature, there is another, related, equation, called the backwards equation, which also describes the flux of neutrons in fissile systems, and is given by

$$\begin{aligned} \frac{\partial}{\partial t} \psi_t(r, v) &= v \cdot \nabla \psi_t(r, v) - \sigma(r, v) \psi_t(r, v) \\ &+ \sigma_{\mathbf{s}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{s}}(r, v, v') dv' + \sigma_{\mathbf{f}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{f}}(r, v, v') dv', \end{aligned} \quad (1.3)$$

with additional boundary conditions

$$\begin{cases} \psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \psi_t(r, v) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } v \cdot \mathbf{n}_r > 0. \end{cases} \quad (1.4)$$

The backwards equation is also known as the adjoint equation since it is the adjoint (in the sense of operator adjoint) of the forwards equation. It describes where the system could evolve to, given the current configuration. Accordingly, the boundary conditions state that neutrons that hit the boundary of the domain with outgoing velocity are “killed” and do not contribute to the system after this time. We will discuss the relationship between the forwards and backwards equations, and their respective boundary conditions, in more detail in Chapter 2.

1.3 Criticality problems

One of the principal ways in which neutron flux is understood is to look for the leading eigenvalue and associated ground state eigenfunction. From the perspective of nuclear engineers, this corresponds to determining the average growth rate of the number of neutrons in the system, and a “heat map” to describe the reactivity profile throughout the reactor. However, there are a variety of different eigenvalue problems associated with the NTE, each leading to a slightly different measure of criticality.

λ -eigenvalue

The λ -, or time-dependent, eigenvalue problem involves finding the largest λ such that there exists a function φ satisfying

$$\begin{aligned} v \cdot \nabla \varphi(r, v) - \sigma(r, v)\varphi(r, v) + \sigma_s(r, v) \int_V \varphi(r, v') \pi_s(r, v, v') dv' \\ + \sigma_f(r, v) \int_V \varphi(r, v') \pi_f(r, v, v') dv' = \lambda \varphi(r, v). \end{aligned} \quad (1.5)$$

Setting λ to be zero in the above equation would indicate a perfect balance between neutron production and absorption. Hence, λ quantifies the imbalance between these two processes, which yields one notion of criticality. In this situation, we have the following regimes:

$$\lambda > 0 \Leftrightarrow \text{supercritical,}$$

$$\lambda = 0 \Leftrightarrow \text{critical,}$$

$$\lambda < 0 \Leftrightarrow \text{subcritical.}$$

Moreover, the eigenfunction φ , also called the *importance map* in the physics literature, gives a profile of the “hot spots” in the reactor. In other words, it describes where the fission activity is most prevalent.

This eigenvalue problem has been considered predominantly in the mathematics literature from a number of perspectives. For example, the works [15, 11] consider this eigenvalue problem from an analytical perspective, while a probabilistic point of view is taken in [16, 12].

k -eigenvalue

The k -eigenvalue problem is a time-independent, or stationary, problem. It corresponds to finding a value $k > 0$ (commonly referred to as k_{eff}) and a function φ_{eff} such that

$$\begin{aligned} v \cdot \nabla \varphi_{\text{eff}}(r, v) - \sigma(r, v) \varphi_{\text{eff}}(r, v) + \sigma_{\text{s}}(r, v) \int_V \varphi_{\text{eff}}(r, v') \pi_{\text{s}}(r, v, v') dv' \\ = -\frac{\sigma_{\text{f}}(r, v)}{k} \int_V \varphi_{\text{eff}}(r, v') \pi_{\text{f}}(r, v, v') dv'. \end{aligned} \quad (1.6)$$

In this case, $k = 1$ corresponds to a perfectly balanced system, and in fact, it can be shown that this occurs precisely when $\lambda = 0$ in the time-dependent problem. In this case, the eigenfunctions φ and φ_{eff} are equal. However, more generally, k can be interpreted as a measure of the average number of neutrons produced from one generation to the next and so, dividing the fission operator by this quantity reestablishes the balance between fission and absorption. Due to this interpretation, k is also used as a measure of criticality, which, in this problem, is categorised as follows:

$$\begin{aligned} k > 1 &\Leftrightarrow \text{supercritical,} \\ k = 1 &\Leftrightarrow \text{critical,} \\ k < 1 &\Leftrightarrow \text{subcritical.} \end{aligned}$$

This eigenvalue problem is the main focus of criticality calculations in industry [3], and has received much attention from the numerical analysis community [10, 17].

It is worth mentioning that this categorisation of criticality is very basic. There is a finer level of classification that takes delayed neutrons into account. For this, we introduce two more parameters, β_{eff} , which is the average number of delayed neutrons produced per fission event, and $\rho = (k - 1)/k$, a measure of reactivity of the reactor. We then have the following regimes:

- **Prompt critical:** $k > 1$; $\rho > \beta_{\text{eff}}$.

The average number of prompt neutrons produced is greater than one and the reactivity of the reactor is higher than the average number of delayed neutrons produced at each fission event. In this case, the number of neutrons is increasing on the same time scale as the production of prompt neutrons. Since this happens very quickly, a system in this regime is highly unstable. For example, this is the situation one would need in order to create a nuclear bomb.

- **Prompt subcritical, delayed supercritical:** $k > 1$; $0 < \rho < \beta_{\text{eff}}$.

Again, the average number of prompt neutrons produced is greater than one, however, in this case it is not these types of neutrons that cause the supercritical state. Of course, the average number of neutrons is still increasing but now on the time scale of delayed neutron production, and it is these latter neutrons that sustain the nuclear reactor and maintain the supercritical state.

- **Prompt subcritical, delayed critical:** $k = 1$; $\rho = 0$.

This is the state in which reactors are designed to operate in and corresponds to the critical state defined above. Indeed, we have $\rho = 0$, and so neutron production and loss are in equilibrium. As in the previous case, the prompt neutrons alone are not enough to sustain the reactivity, and so the delayed neutrons help to achieve this.

- **Prompt subcritical, delayed subcritical:** $k < 1$; $\rho < 0$.

This is where neutron loss outweighs neutron production and corresponds to the subcritical regime above.

c -eigenvalue

This final eigenvalue problem is very similar to the previous one, in the sense that it is a time-independent problem, however there is a subtle difference that arises from requiring the eigenvalue, c , to also weight the scattering operator, as well as the fission operator. More precisely, the aim is to find $c > 0$ and a corresponding function ϕ such that

$$\begin{aligned} & v \cdot \nabla \phi(r, v) - \sigma(r, v)\phi(r, v) \\ &= -\frac{1}{c} \left(\sigma_s(r, v) \int_V \phi(r, v') \pi_s(r, v, v') dv' + \sigma_f(r, v) \int_V \phi(r, v') \pi_f(r, v, v') dv' \right). \end{aligned} \quad (1.7)$$

Similar heuristics to those given for k give the interpretation that c is the effective number of neutrons produced between generations per collision, where, in this case, a collision is either a scatter or a fission. Criticality is also categorised in the same way as the k -eigenvalue problem, and it can be shown that the critical regime also coincides with the critical regime in the previous two problems, with equality of eigenfunctions. This problem has received the least attention in the literature, however, due to its similarities with the k -eigenvalue problem, many of the methods used for solving the latter can be adapted solving the c -eigenvalue problem.

Clearly, in all of these cases, the optimal regime for a reactor is the critical one, since this pertains to a high enough energy production whilst operating at a safe level. In particular, as mentioned above, reactors are designed to operate in the “prompt subcritical, delayed critical” regime. This is the most stable regime and allows the reactor to be operated on the same time scale as the production of delayed neutrons.

In this thesis, we will predominantly focus on the time-dependent problem (1.5), except for Chapter 5, in which we consider the k -eigenvalue problem (1.6). For a more extensive discussion on the various eigenvalue problems, see [17, 10, 13, 2].

1.4 Monte Carlo methods

Although it is not the focus of this thesis, the results obtained in later chapters have allowed us to develop Monte Carlo algorithms that complement existing ones in industry, in order to compute quantities such as the eigenvalues and eigenfunctions discussed in the previous section.

We therefore spend a little time discussing some of the algorithms we have developed in parallel to the theoretical results. These algorithms have been developed further in [6].

The main tool for studying the NTE in this thesis is a branching process, denoted $(X_t)_{t \geq 0}$, whose behaviour corresponds to that of a nuclear fission process. Let $\mathbb{P}_{\delta_{(r,v)}}$ denote its law when issued from a single neutron with configuration (r, v) . Note that for the purposes of this discussion, it is not necessary to give the details of this branching process here; they will be stated in later chapters. By considering an appropriate average of this branching process, we obtain a stochastic analogue of the flux, ψ_t . More precisely, the linear semigroup of this branching process

$$\psi_t[g](r, v) := \mathbb{E}_{\delta_{(r,v)}}[\langle g, X_t \rangle] := \mathbb{E}_{\delta_{(r,v)}} \left[\sum_{i=1}^{N_t} g(r_i(t), v_i(t)) \right],$$

where N_t is the number of neutrons alive at time t and $\{(r_i(t), v_i(t)) : i = 1, \dots, N_t\}$ are their configurations, solves the (backwards) NTE in some sense, and hence, we can think of it as a representation of the flux, ψ_t , with $\psi_0 = g$.

One of the main results we present in this thesis is a Perron-Frobenius decomposition of $\psi_t[g]$, which roughly says that for $(r, v) \in D \times V$,

$$\psi_t[g](r, v) \sim e^{\lambda t} \langle \tilde{\varphi}, g \rangle \varphi(r, v) + o(e^{\lambda t}), \quad t \rightarrow \infty \quad (1.8)$$

where λ , φ and $\tilde{\varphi}$ are the leading eigenvalue, and right and left eigenfunctions, respectively, associated with the λ -eigenvalue problem described in the previous section. Manipulating this asymptotic leads to

$$\lambda = \lim_{t \rightarrow \infty} \frac{1}{t} \log \psi_t[g](r, v). \quad (1.9)$$

In particular, setting $g = \mathbf{1}$, the constant function with value one, we have

$$\lambda = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_{\delta_{(r,v)}}[N_t],$$

which leads to the following Monte Carlo algorithm.

Algorithm 1

Input: $T \geq 0$, $N \in \mathbb{N}$, $(r_0, v_0) \in D \times V$.

for $i = 1, \dots, N$:

1. Simulate a copy of the branching process, $(X_t^{(i)})_{t \geq 0}$ initiated from (r_0, v_0) .
2. Calculate the number of particles, $N_T^{(i)}$, alive at time T .

Output: $\hat{\lambda} = \frac{1}{T} \log \frac{1}{N} \sum_{i=1}^N N_T^{(i)}$.

For large T and N , this algorithm approximates the eigenvalue λ .

We would also like to produce estimates of the left and right eigenfunctions, $\tilde{\varphi}$ and φ . For

this, we define the discounted occupation measure of the branching process up to time t by

$$A_t = \int_0^t e^{-\lambda s} \langle g, X_s \rangle ds.$$

For non-negative and measurable functions g , we can apply Fubini's theorem to switch the expectation and the integral. Then using (1.8), we have

$$\lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}_{\delta_{(r,v)}}[A_t] = \langle \tilde{\varphi}, g \rangle \varphi(r, v), \quad (1.10)$$

which inspires a second algorithm.

Algorithm 2

Input: $g \in L_\infty^+(D \times V)$, $(r_0, v_0) \in D \times V$, $T \geq 0$, $N, M \in \mathbb{N}$

for $i = 1, \dots, N$:

1. Simulate a copy of the branching process, $(X_t^{(i)})_{t \geq 0}$ initiated from (r_0, v_0) .
2. **for** $m = 1, \dots, M$: calculate

$$\langle g, X_{tm/M}^{(i)} \rangle = \sum_{j=1}^{N_{tm/M}^{(i)}} g(r_j^{(i)}(tm/M), v_j^{(i)}(tm/M)),$$

where $N_{tm/M}^{(i)}$ is the number of neutrons in alive at time tm/M , and $(r_j^{(i)}(tm/M), v_j^{(i)}(tm/M))$, $j = 1, \dots, N_{tm/M}^{(i)}$ are their configurations.

Output: $\frac{1}{T} \frac{1}{M} \sum_{m=1}^M \frac{1}{N} \sum_{i=1}^N \langle g, X_{tm/M}^{(i)} \rangle$.

Due to (1.10), for sufficiently large T, M and N , the output of this algorithm gives an estimate of $\langle g, \tilde{\varphi} \rangle \varphi_\lambda(r_0, v_0)$.

Now, on the one hand, fixing g and varying the initial configuration (r_0, v_0) allows us to build a profile of the right eigenfunction φ , up to a multiplicative constant. On the other hand, fixing (r_0, v_0) and varying g allows us to estimate the left eigenfunction $\tilde{\varphi}$. For example, choosing $g = \mathbf{1}_{B(r, \varepsilon)}$, the indicator function of the ball centred at $r \in D$ with radius $\varepsilon > 0$, one can approximate $\tilde{\varphi}$ in a neighbourhood of r .

In short, one can run **Algorithm 2** in an outer loop that iterates over either the test function or the initial condition depending on whether one wants to estimate the left or the right eigenfunction. As an example of this algorithm, we built a toy two-dimensional reactor consisting of a square domain and four uranium ‘‘rods’’. The scattering and fission rates within these rods are much higher than in the rest of the reactor to demonstrate inhomogeneities. The figures below show simulations of the eigenfunctions.

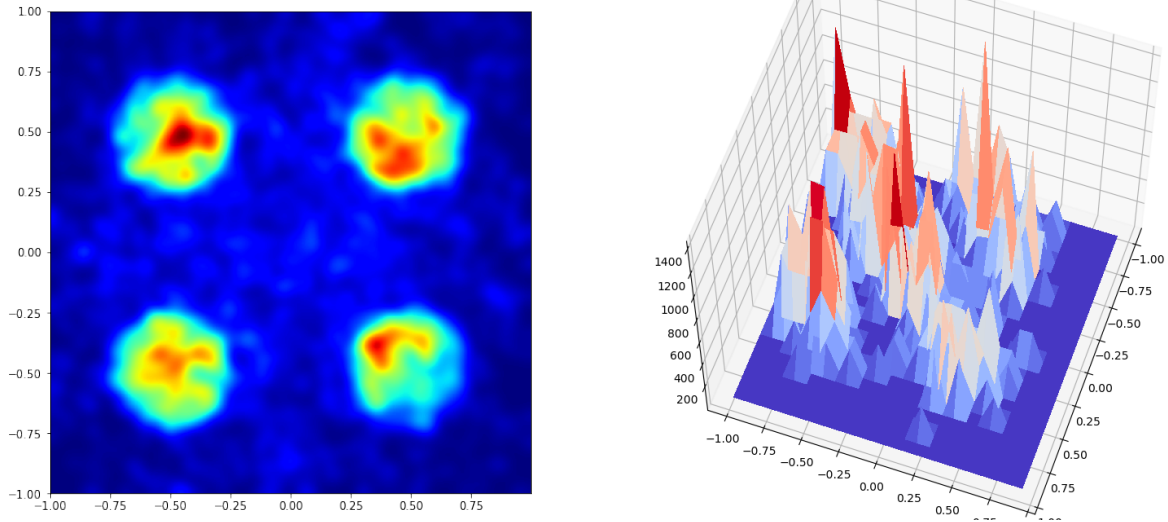


Figure 1-1: A simulation of the eigenfunctions for a two-dimensional reactor containing four uranium rods. The image on the left is a simulation of the left eigenfunction, $\tilde{\varphi}$, and the one on the right is the right eigenfunction, φ .

We note that similar algorithms can be built for estimating the quantities in the other eigenvalue problems. For example, we refer the reader to Chapter 5 for a discussion regarding the k -eigenvalue problem.

In terms of efficiency of these algorithms, the variance associated with the estimators of the eigenvalues is very low since they essentially simulate reality. However, since these algorithms require the simulation of a whole tree of neutrons, the process can become expensive due to increasing particle numbers in even a slightly supercritical regime.

As previously mentioned, there are already several Monte Carlo algorithms in industry for calculating various quantities associated with neutron transport criticality problems. For an overview of some of these methods, we refer the reader to [3, 17, 14, 1, 7].

1.5 Notation

Here we include tables of frequently used abbreviations for the convenience of the reader. We also note that there is a glossary at the end of Chapters 3 and 4 to make it easier for the reader to keep track of commonly used notation.

We also define the following operators for the purpose of the next section. However, they

Abbreviation	Description	Introduced
NTE	Neutron Transport Equation	§1.2
NBP	Neutron Branching Process	§4.1.2
NRW	Neutron Random Walk	§3.6
MNTE	Multi-species Neutron Transport Equation	§2.5
MNBP	Multi-species Neutron Branching Process	§2.6
MNRW	Multi-species Neutron Random Walk	§2.7
NGP	Neutron Generational Populations	§5.2.4

will be restated in the remaining chapters.

$$\left\{ \begin{array}{ll}
\overleftarrow{\mathbb{T}}f(r, v) & := v \cdot \nabla f(r, v) & \text{(backwards transport)} \\
\overleftarrow{\mathbb{S}}f(r, v) & := \sigma_{\mathfrak{s}}(r, v) \int_V f(r, v') \pi_{\mathfrak{s}}(r, v, v') dv' - \sigma_{\mathfrak{s}}(r, v) f(r, v) & \text{(backwards scattering)} \\
\overleftarrow{\mathbb{F}}f(r, v) & := \sigma_{\mathfrak{f}}(r, v) \int_V f(r, v') \pi_{\mathfrak{f}}(r, v, v') dv' - \sigma_{\mathfrak{f}}(r, v) f(r, v) & \text{(backwards fission)}
\end{array} \right. \quad (1.11)$$

1.6 Outline of thesis

The aim of this thesis is to provide a deeper understanding of the neutron transport equation from a probabilistic point of view. Each of these chapters contain a research article, which were written in collaboration with my supervisors Dr Alexander M. G. Cox, Professor Simon C. Harris and Professor Andreas E. Kyprianou, and collaborator Dr Denis Villemonais.

1. **Multi-species neutron transport equation (Chapter 2).** We begin this thesis with the analysis of a multi-type version of the neutron transport equation. As hinted in the introduction, nuclear fission does not only involve neutrons but a whole range of emissions, such as alpha, beta and gamma radiation. From an analytical perspective, this leads to the analysis of a system of equations of the form

$$\frac{\partial \psi_t}{\partial t}(i, r, v) = (\overleftarrow{\mathbb{T}}_i + \overleftarrow{\mathbb{S}}_i + \overleftarrow{\mathbb{F}}_i) \psi_t(i, r, v), \quad t \geq 0, i \in \{1, \dots, m\}, r \in D, v \in V, \quad (1.12)$$

where the index i denotes the particle type, and the operators $\overleftarrow{\mathbb{T}}_i$, $\overleftarrow{\mathbb{S}}_i$ and $\overleftarrow{\mathbb{F}}_i$ are of a similar form to those defined in (1.11). When combined with initial conditions, classical solutions to this problem take the form $u_t \in L^2$ whose time derivative also lies in L^2 , for some appropriate L^2 space.

On the other hand, from a probabilistic perspective, one can build a branching process $X = (X_t)_{t \geq 0}$, whose dynamics match those of the (multi-type) fission process and whose linear semigroup is the stochastic analogue of the flux, ψ_t . More precisely, we can repre-

sent the system by a collection of particles $\{(r_{i,j}(t), v_{i,j}(t)), i = 1, \dots, N_t^j, t \geq 0\}$, where $(r_{i,j}(t), v_{i,j}(t))$ is the space-velocity configuration of the i^{th} type j particle alive at time t . Then the branching process can be defined as the empirical distribution of these particles:

$$X_t(j, A) = \sum_{i=1}^{N_t^j} \delta_{(r_{i,j}(t), v_{i,j}(t))}(A), \quad A \in \mathcal{B}(D \times V), \quad (1.13)$$

and the expectation semigroup is given by

$$\psi_t[g](i, r, v) = \mathbb{E}_{\delta_{(i,r,v)}}[\langle g, X_t \rangle],$$

where $X_t(A) = (X_t(1, A), \dots, X_t(m, A))$.

We would like to be able to match these two perspectives by saying that $\psi_t[g](i, \cdot, \cdot)$ solves (1.12), however the notion of solution to (1.12) is too strong for this to be the case. In the first part of this chapter, we spend time defining an appropriate notion of solution to the NTE in order for us to say that the expectation semigroup does indeed solve (1.12). We then consolidate the two approaches.

The second part of the chapter is devoted to the time-dependent eigenvalue problem (1.5). We use spectral theory to prove the existence of a leading eigentriple $(\lambda, \varphi, \tilde{\varphi})$ and an $\varepsilon > 0$ such that

$$\psi_t[g](r, v) = e^{\lambda t} \langle g, \tilde{\varphi} \rangle \varphi(r, v) + \mathcal{O}(e^{(\lambda-\varepsilon)t}), \quad t \rightarrow \infty, \quad (1.14)$$

where $\psi_t[g](r, v) = (\psi_t[g](1, r, v), \dots, \psi_t[g](m, r, v))$. In other words, the eigentriple characterises the leading order behaviour of the branching system.

Cox, A.M.G. and Harris, S.C. and Horton, E. and Kyprianou, A.E. (2019) Multispecies Neutron Transport Equation, Journal of Statistical Physics

- 2. Linear semigroup asymptotics (Chapter 3).** We now restrict ourselves to the basic NTE where only (prompt) neutrons are considered. Although we were already able to provide a solution to the criticality problem (1.5) in Chapter 2, this is still unsatisfactory from a probabilistic point of view, since we were forced to work in L^2 spaces, whereas the more natural setting for expectation semigroups is L^∞ and its dual, L^1 .

In this chapter we develop a probabilistic approach to characterising the growth of $\psi_t[g]$. The main tools for this approach come from the theory of quasi-stationary distributions that were developed in [4] and [5]. The first step is to show that the neutron branching process is equivalent to another process, which we call the neutron random walk, via a many-to-one formula. Indeed, one can rearrange the NTE to obtain

$$\tilde{\mathbf{T}} + \tilde{\mathbf{S}} + \tilde{\mathbf{F}} = \tilde{\mathbf{L}} + \beta,$$

where $\tilde{\mathbf{L}}$ is an operator that has the same form as $\tilde{\mathbf{T}} + \tilde{\mathbf{S}}$ but with a different rate and

kernel, and β is a local potential. This inspires the Feynman-Kac representation

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

where $(R_s, \Upsilon_s)_{s \geq 0}$ is the process with generator $\overleftarrow{\mathbf{L}}$. Under certain assumptions, one can consider a sub-Markov process that also has generator $\overleftarrow{\mathbf{L}}$ but is eventually absorbed. Studying the quasi-stationary distribution of this latter process allows us to characterise its growth in a similar way to (1.14), which immediately gives us the leading order behaviour of the semigroup ψ_t .

In the rest of this chapter we consider the martingale

$$W_t = e^{-\lambda_* t} \frac{\langle \varphi, X_t \rangle}{\langle \varphi, \mu \rangle}, \quad (1.15)$$

where (λ_*, φ) are the leading eigenvalue and right eigenfunction of ψ_t , and $\mu = \sum_{i=1}^n \delta_{(r_i, v_i)}$ describes some initial configuration of particles. We show that there are two regimes for the long-term behaviour of $(W_t)_{t \geq 0}$, depending on whether $\lambda_* > 0$ or $\lambda_* \leq 0$. Finally, we use $(W_t)_{t \geq 0}$ to define a change of measure, which leads to a spine decomposition for the neutron branching process.

Preprint: arXiv:1810.01779

3. **Skeleton decompositions and the strong law of large numbers (Chapter 4).** We revisit the basic NTE from the previous chapter but instead focus on the supercritical case, $\lambda_* > 0$. Our main goal in this chapter is to characterise the growth of the system via a strong law of large numbers (SLLN) result:

$$\lim_{t \rightarrow \infty} e^{-\lambda_* t} \frac{\langle g, X_t \rangle}{\langle \varphi, \mu \rangle} = \langle g, \tilde{\varphi} \rangle W_\infty, \quad \text{almost surely,}$$

where $(\lambda_*, \varphi, \tilde{\varphi})$ is the leading eigentriple from the previous chapter, and W_∞ is the limit of the martingale $(W_t)_{t \geq 0}$, defined in (1.15).

Our methods for proving this result are inspired by those presented in [8]. Indeed, we would like to be able to adapt the authors' proofs directly to the neutron branching process. However, their proof relies on the fact that the number of particles in the system is non-decreasing, whereas in our setting, neutrons can be absorbed. This leads us to consider a skeleton decomposition, similar in spirit to [9]. The idea is to split the neutron branching process into a tree, X^\uparrow , which contains all the particles that survive forever, and subtrees X^\downarrow , which are attached to X^\uparrow and all become extinct. We are then in a position to apply the proof of [8] to X^\uparrow to obtain a SLLN for this process. It then remains to show that the subtrees X^\downarrow contribute nothing to the limit, due to the fact that they eventually become extinct.

Preprint: arXiv:1901.00220

4. **Time-independent criticality problems (Chapter 5).** In this final chapter, we move away from the time-dependent eigenvalue problem and consider the stationary problems. In particular, we focus on the k -eigenvalue problem, since this is of particular interest in industry; however our methods will also apply with only minor modifications to the c -eigenvalue problem.

The k -eigenvalue problem pertains to finding a positive parameter k and associated eigenfunction φ_{eff} such that

$$(\bar{\mathbf{T}} + \bar{\mathbf{S}} - \sigma_{\mathbf{f}}\mathbf{I})\varphi = -\frac{1}{k}(\bar{\mathbf{F}} + \sigma_{\mathbf{f}}\mathbf{I})\varphi,$$

so that k can be interpreted as the average number of neutrons produced per fission event.

Again, we would like to view this problem from both an analytical and probabilistic point of view, and consolidate the two approaches. In order to address the existence of $(k, \varphi_{\text{eff}})$ in the classical sense, we apply similar techniques to those employed in Chapter 2. For the probabilistic approach, it turns out that the correct object to study is the semigroup Ψ_n associated with the average of the neutron branching process at the n^{th} generation of fission events. As in Chapter 3, we construct a many-to-one formula in order to study the quasi-stationary distribution of an associated killed random walk in order to obtain $(k, \varphi_{\text{eff}}, \tilde{\varphi}_{\text{eff}})$ and $\delta > 1$ such that

$$\Psi_n[g](r, v) = k^n \langle g, \tilde{\varphi}_{\text{eff}} \rangle \varphi_{\text{eff}}(r, v) + \mathcal{O}(\delta^{-n}), \quad n \rightarrow \infty.$$

Preprint: arXiv:1909.00581

This thesis is presented in the alternative format which includes publications. This means the research chapters are developed independent of the introduction and supposed to be self-contained. Hence, it is inevitable that there will be some inconsistencies in notation and redundant content in the introduction chapter.

Bibliography

- [1] G. I. Bell and S. Glasstone. *Nuclear Reactor Theory*. Reinhold, 1970.
- [2] B. R. Betzler. Calculating alpha eigenvalues and eigenfunctions with a Markov transition rate matrix Monte Carlo method. *PhD Thesis*, 2014.
- [3] R. J. Brissenden and Garlick A.R. Biases in the estimation of k_{eff} and its error by Monte Carlo methods. *Ann. nucl. Energy*, 13(2):63–83, 1986.
- [4] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q -process. *Probab. Theory Related Fields*, 164(1-2):243–283, 2016.
- [5] N. Champagnat and D. Villemonais. Uniform convergence to the Q -process. *Electron. Commun. Probab.*, 22:Paper No. 33, 7, 2017.
- [6] A. M. G. Cox, S.C. Harris, A.E. Kyprianou, and M. Wang. Monte Carlo methods for the neutron transport equation. *Working document*, 2018.
- [7] B. Davison and J. B. Sykes. *Neutron transport theory*. Oxford, at the Clarendon Press, 1957.
- [8] J. Engländer, S. C. Harris, and A. E. Kyprianou. Strong law of large numbers for branching diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):279–298, 2010.
- [9] S. C. Harris, M. Hesse, and A. E. Kyprianou. Branching Brownian motion in a strip: survival near criticality. *Ann. Probab.*, 44(1):235–275, 2016.
- [10] M.J. Parkinson I.G. Graham and R. Scheichl. Modern Monte Carlo variants for uncertainty quantification in neutron transport. *Festschrift for the 80th Birthday of Ian Sloan, eds. J. Dick, F.Y. Kuo, and H. Wozniakowski*, 2018.
- [11] K. Jörgens. An asymptotic expansion in the theory of neutron transport. *Comm. Pure Appl. Math.*, 11:219–242, 1958.
- [12] B. Lapeyre, É. Pardoux, and R. Sentis. *Introduction to Monte-Carlo methods for transport and diffusion equations*, volume 6 of *Oxford Texts in Applied and Engineering Mathematics*. Oxford University Press, Oxford, 2003. Translated from the 1998 French original by Alan Craig and Fionn Craig.

- [13] E. E. Lewis and Jr. W. F. Miller. *Computational Methods of Neutron Transport*. Wiley & Sons, New York, 1984.
- [14] S. Maire and D. Talay. On a Monte Carlo method for neutron transport criticality computations. *IMA J. Numer. Anal.*, 26(4):657–685, 2006.
- [15] M. Mokhtar-Kharroubi. *Mathematical topics in neutron transport theory*, volume 46 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co., Inc., River Edge, NJ, 1997. New aspects, With a chapter by M. Choulli and P. Stefanov.
- [16] T. Mori, S. Watanabe, and T. Yamada. On neutron branching processes. *Publ. Res. Inst. Math. Sci.*, 7:153–179, 1971/72.
- [17] F. Scheben and I. G. Graham. Iterative methods for neutron transport eigenvalue problems. *SIAM J. Sci. Comput.*, 33(5):2785–2804, 2011.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Multi-species Neutron Transport Equation			
Publication status (tick one)			
Draft manuscript	<input type="checkbox"/>	Submitted	<input type="checkbox"/>
		In review	<input type="checkbox"/>
		Accepted	<input type="checkbox"/>
		Published	<input checked="" type="checkbox"/>
Publication details (reference)	doi.org/10.1007/s10955-019-02307-2		
Copyright status (tick the appropriate statement)			
I hold the copyright for this material		<input type="checkbox"/>	Copyright is retained by the publisher, but I have been given permission to replicate the material here
			<input checked="" type="checkbox"/>
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	<p>*All research has been conducted in equal partnership with collaborators and supervisors. It is unwise to try and measure percentages of intellectual contribution as research scholarship lies as much in the escalation of ideas through mathematical discourse as it does with the seed of ideas themselves. That said, the mathematical content of this thesis is, as an entire piece of work, inextricably associated to its author through intellectual ownership.</p> <p>Formulation of ideas:</p> <p>The candidate played an integral and fully collaborative role in the formulation of ideas.</p> <p>Design of methodology:</p> <p>The candidate played an integral and fully collaborative role in the design of methodology.</p> <p>Experimental work:</p> <p>N/A</p> <p>Presentation of data in journal format:</p> <p>N/A</p> <p>*The wording in this box follows the advice and approval of my supervisor, Professor Kyprianou.</p>		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Emma Horton	Date	19/11/2019

Chapter 2

Multi-species neutron transport equation

Alex M. G. Cox¹, Simon C. Harris², Emma Horton³, Andreas E. Kyprianou⁴.

Abstract

The Neutron Transport Equation (NTE) describes the flux of neutrons through an inhomogeneous fissile medium. Whilst well treated in the nuclear physics literature (cf. [8, 29]), the NTE has had a somewhat scattered treatment in mathematical literature with a variety of different approaches (cf. [7, 27]). Within a probabilistic framework it has somewhat undeservingly received little attention in recent years; nonetheless, probabilistic treatments can be found: see for example [19, 28, 24, 31, 4, 3]. In this article our aim is threefold. First we want to introduce a slightly more general setting for the NTE, which gives a more complete picture of the different species of particle and radioactive fluxes that are involved in fission. Second we consolidate the classical c_0 -semigroup approach to solving the NTE with the method of stochastic representation which involves expectation semigroups. Third we provide the leading asymptotic of our multi-species NTE, which will turn out to be crucial for further stochastic analysis of the NTE in forthcoming work [15, 13, 5]. The methodology used in this paper harmonises the culture of expectation semigroup analysis from the theory of stochastic processes against c_0 -semigroup theory from functional analysis. In this respect, our presentation is thus part review of existing theory and part presentation of new research results based on generalisation of existing results.

¹Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: a.m.g.cox@bath.ac.uk

²Department of Statistics, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. Email: simon.harris@auckland.ac.nz

³Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: elh48@bath.ac.uk

⁴Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: a.kyprianou@bath.ac.uk

2.1 Introduction

The neutron transport equation (NTE) describes the flux of neutrons across a directional planar cross-section in an inhomogeneous fissile medium (typically measured in number of neutrons per cm^2 per second). As such, the flux is described as a function of time, t , Euclidian location, $r \in \mathbb{R}^3$, direction of travel, $\Omega \in \mathbb{S}_2$, and neutron energy, $E \in \mathbb{R}$. It is not uncommon in the physics literature, as indeed we shall do here, to assume that energy is a function of velocity ($E = m|v|^2/2$), thereby reducing the number of variables by one. This allows us to describe the dependency of flux more simply in terms of time and, what we call, the *configuration variables* $(r, v) \in D \times V$ where $D \subseteq \mathbb{R}^3$ is a smooth, open, connected and bounded domain of concern such that ∂D has zero Lebesgue measure and V is the velocity space, which can now be taken to be $V = \{v \in \mathbb{R}^3 : v_{\min} < |v| < v_{\max}\}$, where $0 < v_{\min} < v_{\max} < \infty$ ⁵.

Before stating the NTE, let us remind the reader of some elementary nuclear physics, which is required to describe the evolution of neutron flux. In the most basic of flux models, there are essentially only four processes at the level of the atomic nuclei which contribute to the evolution of neutron flux.

The first is *spontaneous neutron emission* from unstable nuclei. This comes from radioactive isotopes whose nuclei are excited. They cause what is known as *non-transmutation emissions*, in which a neutron is ejected with an escape velocity (*neutron emission*), or, conversely, what are called *transmutation emissions* in which the nucleus instantaneously fragments into two or more nuclei (*spontaneous fission*) with a range of possible masses, emitting one or more neutrons with escape velocities in the process.

The second process pertains to *neutron scattering*. This is where a neutron travelling with a given velocity passes in close proximity to an atomic nucleus, which, in our model, results in an instantaneous change of velocity.

The third process is *neutron-induced fission*. This is the classical setting in which a neutron travelling with a given velocity strikes an atomic nucleus sending it into an excited state, from which it instantaneously fragments into two or more nuclei, simultaneously releasing one or more neutrons.

The fourth and final process is *neutron capture*. In this setting, a neutron travelling with a given velocity strikes an atomic nucleus, but instead of causing nuclear fission, it is absorbed into the nucleus. It can also be the case that neutrons decay into other subatomic particles, and thus disappear from the system. To all intents and purposes, we can treat this as neutron capture.

When modelling the transmission of neutrons in a fissile material, those neutrons which have been released from nuclei are known as *prompt neutrons*.

With more advanced modelling, one can also take account of the fact that some of the processes described above can also involve other types of nuclear emissions, often in addition to neutrons. These include alpha and beta particles and gamma radiation. Whilst the former

⁵In practice, $v_{\min} = 1/40\text{MeV}$ and v_{\max} can be taken to be the speed of light.

two are not sufficiently energetic to cause fission, sufficiently energetic gamma rays are able to induce fission.

Spontaneous fission and neutron-induced fission can also produce what are known as *delayed neutrons*. These are neutrons released from a fission product (isotope) some time after fission has occurred. In terms of modelling, they are spontaneous neutron emissions which occur at the site of neutron-induced fission but at a moment later in time. Delayed neutrons are only in a delayed state until they are released after which they are considered as prompt neutrons.

We refer to models which take account of the full range of flux profiles as multi-species models.

2.2 Neutron Transport Equation

Let us now write down the basic neutron transport equation (prompt neutrons only), which has been widely considered in a variety of physics and engineering literature (cf. [8, 29], to name but two classical references), and somewhat more sporadically studied in the mathematical literature. See [7, 27, 19] for the three most authoritative mathematical texts in more recent times, as well as e.g. [28, 14, 24] for some of the rarer examples of the probabilistic treatment of the NTE.

Neutron flux at time $t \geq 0$ is henceforth identified as $\Psi_t : D \times V \rightarrow [0, \infty)$, and the classical presentation of its evolution in time is given by the integro-differential equation, also known as the *forward neutron transport equation*⁶,

$$\begin{aligned} \frac{\partial}{\partial t} \Psi_t(r, v) &= -v \cdot \nabla \Psi_t(r, v) - \sigma(r, v) \Psi_t(r, v) + Q(r, v, t) \\ &+ \int_V \Psi_t(r, v') \sigma_s(r, v') \pi_s(r, v', v) dv' + \int_V \Psi_t(r, v') \sigma_f(r, v') \pi_f(r, v', v) dv', \end{aligned} \quad (2.1)$$

where the different components (or *cross-sections as they are known in the physics literature*) are all uniformly bounded and measurable with the following interpretation:

- $\sigma_s(r, v')$: the rate at which scattering occurs from incoming velocity v' ,
- $\sigma_f(r, v')$: the rate at which fission occurs from incoming velocity v' ,
- $\sigma(r, v)$: the sum of the rates $\sigma_f + \sigma_s$, also known as the *total* cross section
- $\pi_s(r, v', v) dv'$: the scattering yield at velocity v from incoming velocity v' ,
satisfying $\int_V \pi_s(r, v, v') dv' = 1$,
- $\pi_f(r, v', v) dv'$: the neutron yield at velocity v from fission with incoming velocity v' ,
satisfying $\int_V \pi_f(r, v, v') dv' < \infty$, and
- $Q(r, v, t)$: non-negative source term.

It is normal to assume that all quantities are uniformly bounded away from infinity. It is also

⁶Here and everywhere else in the document, ∇ is the gradient operator with respect to the variable $r \in \mathbb{R}^3$.

usual to assume the additional boundary conditions

$$\begin{cases} \Psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \Psi_t(r, v) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } v \cdot \mathbf{n}_r < 0, \end{cases} \quad (2.2)$$

where \mathbf{n}_r is the outward facing normal of D at $r \in \partial D$ and $g : D \times V \rightarrow [0, \infty)$ is a bounded, measurable function which we will later assume has some additional properties. Roughly speaking, as the forward equation describes where particles could have evolved from in order to contribute to the current configuration, this boundary condition means that particles from outside the domain with incoming velocity are not taken into account. The second of the above two boundary conditions is sometimes written $\Psi_t|_{\partial D^-} = 0$, where $\partial D^- = \{(r, v) \in \partial D \times V : v \cdot \mathbf{n}_r < 0\}$. It is also usual to set $Q = 0$ when considering a reactor with a multiplying medium, as the resulting fission will overwhelm the radioactive source term.

The notion of a solution of the form (2.1) turns out to be too strong to expect to make mathematical sense of it. This is predominantly due to the non-diffusive nature of the equation, in particular the non-local nature of the scattering and fission operators as well as regularity issues on the domain $D \times V$ in relation to continuity properties of e.g. the operator $v \cdot \nabla$. It is much more natural to look for solutions that belong to e.g. an appropriate L_2 space. This is, moreover, helpful when looking to understand (2.1) as a *backwards* equation, rather than a *forwards* equation.

With some rearrangements, the components of (2.1) separate into transport, scattering and fission. Specifically,

$$\begin{cases} \vec{\mathbb{T}}g(r, v) := -v \cdot \nabla g(r, v) - \sigma(r, v)g(r, v) & \text{(forwards transport)} \\ \vec{\mathbb{S}}g(r, v) := \int_V g(r, v') \sigma_{\mathbf{s}}(r, v) \pi_{\mathbf{s}}(r, v', v) dv' & \text{(forwards scattering)} \\ \vec{\mathbb{F}}g(r, v) := \int_V g(r, v') \sigma_{\mathbf{f}}(r, v) \pi_{\mathbf{f}}(r, v', v) dv' & \text{(forwards fission)} \end{cases} \quad (2.3)$$

such that all operators are defined on $D \times V$ and their action is zero otherwise. Let us momentarily consider the operator on the right-hand side of (2.1) as acting on $L_2(D \times V)$, the space of square integrable functions on $D \times V$, and write

$$\langle f, g \rangle = \int_{D \times V} f(r, v) g(r, v) dr dv$$

for the associated inner product. Note that, for $f, g \in L_2(D \times V)$ such that both $v \cdot \nabla f$ and $v \cdot \nabla g$ are well defined as distributional derivatives, which are also in the space $L_2(D \times V)$, with g respecting the second of the boundary conditions in (2.2), we can verify with a simple

integration by parts that, for $v \in V$,

$$\langle f, v \cdot \nabla g \rangle = \int_{\partial D \times V} (v \cdot v') f(r, v') g(r, v') dr dv' - \langle v \cdot \nabla f, g \rangle = -\langle v \cdot \nabla f, g \rangle \quad (2.4)$$

providing we insist that f respects the boundary $f(r, v) = 0$ for $r \in \partial D$ if $v \cdot \mathbf{n}_r > 0$. Moreover, Fubini's theorem also tells us that, for example, with $f, g \in L_2(D \times V)$,

$$\begin{aligned} \langle f, \int_V g(\cdot, v') \sigma_{\mathbf{s}}(\cdot, v') \pi_{\mathbf{s}}(\cdot, v', \cdot) dv' \rangle &= \int_{D \times V \times V} f(r, v) \sigma_{\mathbf{s}}(r, v') g(r, v') \pi_{\mathbf{s}}(r, v', v) dv' dr dv \\ &= \int_{D \times V} \sigma_{\mathbf{s}}(r, v') \int_V f(r, v) \pi_{\mathbf{s}}(r, v', v) dv g(r, v') dr dv' \\ &= \langle \sigma_{\mathbf{s}}(\cdot, \cdot) \int_V f(\cdot, v) \pi_{\mathbf{s}}(\cdot, \cdot, v) dv, g \rangle. \end{aligned}$$

These computations tell us that, for $f, g \in L_2(D \times V)$ such that $v \cdot \nabla g$ and $v \cdot \nabla f$ are well defined in the distributional sense and in L^2 , and moreover, that $g(r, v) = 0$ for $r \in \partial D$ if $v \cdot \mathbf{n}_r < 0$, and for $f \in L_2(D \times V)$ such that $f(r, v) = 0$ for $r \in \partial D$ if $v \cdot \mathbf{n}_r > 0$,

$$\langle f, (\vec{\mathbf{T}} + \vec{\mathbf{S}} + \vec{\mathbf{F}})g \rangle = \langle (\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})f, g \rangle,$$

where now we identify the transport, scattering and fission operators as

$$\left\{ \begin{array}{ll} \overleftarrow{\mathbf{T}} f(r, v) & := v \cdot \nabla f(r, v) & \text{(backwards transport)} \\ \overleftarrow{\mathbf{S}} f(r, v) & := \sigma_{\mathbf{s}}(r, v) \int_V f(r, v') \pi_{\mathbf{s}}(r, v, v') dv' - \sigma_{\mathbf{s}}(r, v) f(r, v) & \text{(backwards scattering)} \\ \overleftarrow{\mathbf{F}} f(r, v) & := \sigma_{\mathbf{f}}(r, v) \int_V f(r, v') \pi_{\mathbf{f}}(r, v, v') dv' - \sigma_{\mathbf{f}}(r, v) f(r, v) & \text{(backwards fission)} \end{array} \right. \quad (2.5)$$

such that all operators are defined on $D \times V$ with zero action otherwise. The reader will immediately note that, although the terms in the sum $\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}$ are identifiable as the adjoint of the terms in the sum $\vec{\mathbf{T}} + \vec{\mathbf{S}} + \vec{\mathbf{F}}$, the same can not be said for the individual 'T', 'S' and 'F' operators. That is to say, the way we have grouped the terms does not allow us to say that e.g. $\overleftarrow{\mathbf{T}}$ is the adjoint operator to $\vec{\mathbf{T}}$ and so on.

The reason for this difference in grouping of terms lies with how one reads the operators in terms of infinitesimal generators as a probabilist. Although this will not make any difference in the analysis of this paper, we keep to this notation for the sake of consistency with further related articles which offer a probabilistic perspective on the backwards NTE; see [5, 15, 13].

Roughly speaking, $\overleftarrow{\mathbf{T}}$, with an appropriately defined domain, is the generator of the rather simple Markov process consisting of a deterministic motion with velocity v , i.e. transport due to pure advection, with killing on exiting the domain D . Similarly, with an appropriately defined domain, the operator $\overleftarrow{\mathbf{S}}$ is the generator corresponding to scattering, in which a particle travelling with velocity v at position r is removed at rate $\sigma_{\mathbf{s}}$ and replaced by a new particle at r with velocity v' chosen with probability $\pi_{\mathbf{s}}(r, v, v') dv'$. Taking advantage of the fact that

$\int_V \pi_{\mathbf{s}}(r, v, dv') dv' = 1$ we can also write

$$\sigma_{\mathbf{s}}(r, v) \int_V f(r, v') \pi_{\mathbf{s}}(r, v, v') dv' - \sigma_{\mathbf{s}}(r, v) f(r, v) = \sigma_{\mathbf{s}}(r, v) \int_V [f(r, v') - f(r, v)] \pi_{\mathbf{s}}(r, v, v') dv'$$

and also note that it takes the classical form of a difference operator. Finally $\overleftarrow{\mathbf{F}}$ is the generator action of a fission even in which a particle travelling with velocity v at position r is removed at rate $\sigma_{\mathbf{f}}$ and replaced by an average number of particles $\pi_{\mathbf{f}}(r, v, v') dv'$ moving onwards from r with velocity v' .

This leads us to the so called *backwards neutron transport equation* (which is also known as the *adjoint neutron transport equation*) given by

$$\begin{aligned} \frac{\partial}{\partial t} \psi_t(r, v) &= v \cdot \nabla \psi_t(r, v) - \sigma(r, v) \psi_t(r, v) \\ &+ \sigma_{\mathbf{s}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{s}}(r, v, v') dv' + \sigma_{\mathbf{f}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{f}}(r, v, v') dv', \end{aligned} \quad (2.6)$$

with additional boundary conditions

$$\begin{cases} \psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \psi_t(r, v) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } v \cdot \mathbf{n}_r > 0. \end{cases} \quad (2.7)$$

Similarly to previously, the second of these two conditions is often written $\psi_t|_{\partial D^+} = 0$, where $\partial D^+ := \{(r, v) \in \partial D \times V : v \cdot \mathbf{n}_r > 0\}$.

The NTE has played a prominent role in real-world modelling and, for many years, has found a home in commercial software⁷ which is used in the nuclear safety industry. In particular, this is most prominent in the modelling and design of environments which are exposed to radioactive material, from nuclear reactor cores and hospital equipment, through to equipment used to irradiate produce that is sold in supermarkets, thereby prolonging its shelf-life. More recently, with the notion of human interplanetary space exploration becoming less of a sci-fi fantasy and more of a fast approaching reality, an understanding of how long-lasting and compact nuclear power sources, for e.g. Moon or Mars bases has become increasingly important.

Figure 2-1 below illustrates a typical geometrical model of a reactor core rod, cladding and outer shielding.⁸ The structural design of such a reactor can easily be stored as virtual environment (i.e. storing the coordinates of the different geometrical domains and the material properties in each domain) with around 150MB of data, on to which extensive data libraries of numerical values for the respective quantities $\sigma_{\mathbf{s}}, \sigma_{\mathbf{f}}, \pi_{\mathbf{s}}, \pi_{\mathbf{f}}$ can be mapped. (It is an otherwise little known fact that countries which are heavily invested in nuclear power, such as the UK, USA, France, China, etc., are all in possession of such numerical libraries of cross sections, which have been carefully built up over decades.)

⁷MONK and MCBEND codes, for example

⁸The authors are grateful to Prof. Paul Smith from Wood who has given us permission to use these images which were constructed with Wood nuclear software ANSWERS.

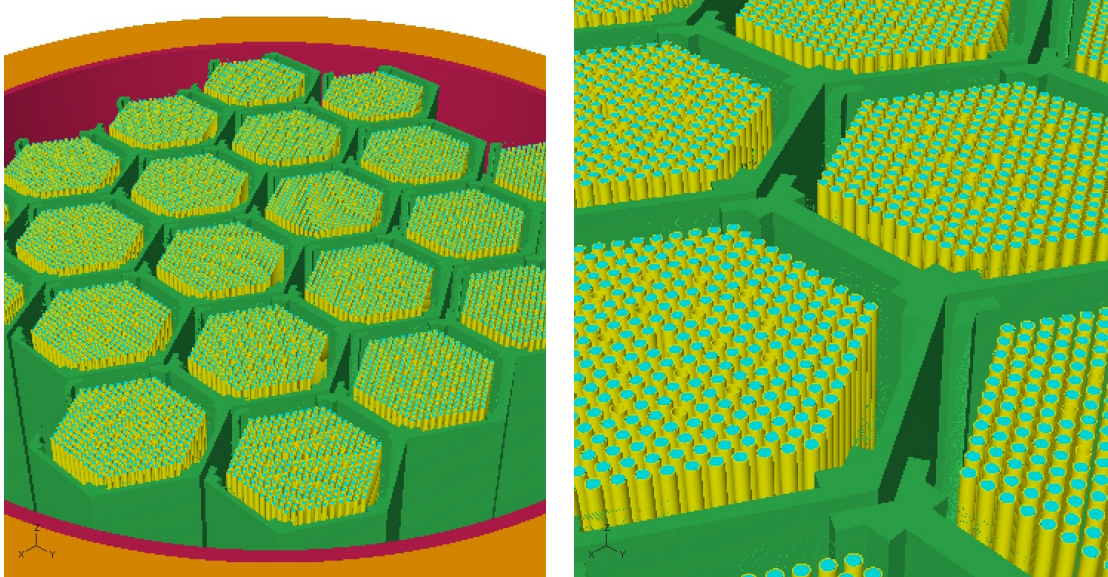


Figure 2-1: A virtual model of a nuclear reactor core with colour indicating the respective fissile properties of the virtual materials used. Uranium rods are arranged into hexagonal cells which are arranged within a larger containment casing.

One of the principal ways in which neutron flux is understood is to look for the leading eigenvalue and associated ground state eigenfunction. Roughly speaking, this means looking for an associated triple of eigenvalue $\lambda \in \mathbb{R}$, non-negative right eigenfunction $\varphi : D \times V \rightarrow [0, \infty)$ in $L_2(D \times V)$ satisfying $\varphi|_{\partial D^+} = 0$ and a non-negative left eigenfunction $\tilde{\varphi}$ on $D \times V$ in $L_2(D \times V)$ satisfying $\tilde{\varphi}|_{\partial D^-} = 0$ such that

$$\lambda \langle \varphi, f \rangle = \langle (\tilde{\mathbf{T}} + \tilde{\mathbf{S}} + \tilde{\mathbf{F}})\varphi, f \rangle \quad \text{and} \quad \lambda \langle f, \tilde{\varphi} \rangle = \langle (\tilde{\mathbf{T}} + \tilde{\mathbf{S}} + \tilde{\mathbf{F}})f, \tilde{\varphi} \rangle.$$

As such, this introduces the notion of *fissile stability*, in particular in the case that $\lambda = 0$. This is naturally the desired scenario⁹ for a nuclear reactor.

In the physics literature, it is thus often understood that, to leading order, the NTE (2.6) is solved in the approximate sense

$$\psi_t(r, v) = e^{\lambda t} \langle g, \tilde{\varphi} \rangle \varphi(r, v) + o(e^{\lambda t}), \quad t \geq 0. \quad (2.8)$$

Note that the scenario that $\lambda > 0$ is obviously to be avoided in practice as this would correspond to a set-up that could result in exponential growth in fission.

The approximation (2.8) can be seen as a functional version of the Perron-Frobenius Theorem and has given rise to a number of different numerical methods for estimating the value of the eigenvalue λ as well as the eigenfunctions φ and $\tilde{\varphi}$. One approach pertains to the discretisation of (2.1) followed by the use of numerical analytic methods; see [32]. Another pertains to the previously alluded to identification of the solution to the NTE as the linear semigroup of a

⁹Strictly speaking the reality is that, nuclear reactors are kept in a slightly supercritical state. The reason for this is that at criticality, as proved in [13], neutron activity will eventually die out.

Markov branching process, which in turn implies Monte Carlo methods involving the simulation of the aforesaid branching process. Such methods are computationally expensive, as branching processes, being tree-like structures, are complex to simulate, e.g. from the point of view of parallelisation. In related papers to this one, we will discuss a new Monte Carlo approach to the NTE based on some of the stochastic analysis we deal with in this article as well as in related work undertaken by the authors of this paper; see [15, 13, 5].

The aim of this paper is as follows. First and foremost, we aim to reposition the theory of the NTE into a contemporary probabilistic setting. We will do this by explaining a precise relationship between the NTE and two different families of Markov processes via Feynman–Kac type formulae. Indeed, this article is one of a cluster of forthcoming pieces of work, which take a new and predominantly probabilistic point of view of the NTE; cf [5, 15, 13]. Next we want to introduce the notion of the (multi-species) NTE into the literature, which generalises (2.1) by simultaneously modeling the flux of all species of particles and radiation involved in the process of nuclear fission. In doing so we will show that, just as in the classical setting, one may develop the notion of a lead eigenvalue and eigenfunction, which is an important part of describing fissile stability. As such, the current article is part review of existing theory and part presentation of new research results based on generalisation of existing results.

Together with the accompanying papers [5, 15, 13], we believe that the probabilistic perspective presented here, i.e. coupling the solutions to the NTE with averaging procedures of certain Markov processes, opens up the possibility of many questions that can be considered at depth in the arena of stochastic analysis and Monte Carlo algorithms, which are currently missing from the literature. Indeed, returning to the kind of environments seen in Figure 2-1, there are many questions concerning how to analyse and numerically generate the leading eigenfunctions and eigenvalue to a reasonable degree of precision. Such questions might include: *What is the connection of the eigendecomposition discussed in this paper and e.g. R-theory or the theory of general Harris recurrence for stochastic processes (cf. [10, 26, 25])? How do different stochastic representations lead to different Monte Carlo simulations? Based on stochastic representation how does one measure convergence of Monte Carlo algorithms? How strong can they be predicted to be? What kind of variance reduction techniques does stochastic representation suggest? Does the inclusion of multi-species models make estimation of the leading eigenvalue more accurate?*

2.3 Organisation of the paper

In the next section, we give a brief overview of the key mathematical literature for the NTE. (Note we do not stray beyond mathematical literature, as the physics and engineering literature is significantly more expansive.) Thereafter in Section 2.5, we introduce the multi-species NTE (MNTE) and its rigorous formulation, existence, uniqueness and asymptotics in the setting of an abstract Cauchy problem. In particular, we show how the unique solution is identified as a c_0 -semigroup in the appropriate L_2 space. In Section 2.6, we introduce a spatial branching process that is constructed using the cross sections that appear in the NTE to describe its stochastic evolution. Here we introduce its expectation semigroup. In Section 2.7, we provide a second

stochastic representation to the expectation semigroup introduced in the previous section via a classical method of the many-to-one formula.

Ideally, we would like to claim that the expectation semigroup discussed in Sections 2.6 and 2.7 agree with the c_0 -semigroup introduced in Section 2.5 (its formal definition appearing just above Theorem 2). This is particularly desired as it forms the foundations of how Monte Carlo simulation of the physical process can be used to develop a numerical solution to the MNTE. In Section 2.8, we consolidate the two notions of semigroup and show that there is partial agreement in an appropriate sense. As far as we are aware, this is a point which is currently not clearly discussed in the literature. Finally we end the paper with a proof of one of the main theorems in Section 2.6 which provides the asymptotic behaviour of the solution to the MNTE in terms of the lead eigenfunction. This is a new result in the multi-species setting in the sense that we have allowed for multiple types of prompt emissions (both particles and radioactive emissions) rather than the case of only one type of prompt emission dealt with in [27]; we also allow for multiple types of delayed emissions (that is, emissions that are pre-emptively held in an unstable radioactive isotope product from an earlier fission event). Our proof nonetheless takes inspiration from the classical approach of [7, 27], and remains loyal to the techniques there.

2.4 Historical remarks on the mathematical treatment of the NTE

Classical texts such as Davison and Sykes [8] were once hailed as a bible of mathematical knowledge during the 1950s post *Manhattan Project* era when rapid technological advances lead to the construction of the very first nuclear reactors driving commercial power stations. Around this time, there was an understanding of how to treat the NTE in special geometries and also by imposing isotropic scattering and fission, see for example Lehner [20] and Lehner and Wing [22, 21]. It was also understood quite early on that the natural way to cite the NTE is via the linear differential transport equation associated to a suitably defined operator on a Banach space. Moreover, it was understood that in this formulation, a spectral decomposition should play a key role in representing solutions, see e.g. Jörgens [17], Pazy and Rabinowitz [30]. This notion was promoted by the work of R. Dautray and collaborators, who showed how c_0 -semigroups form a natural framework within which one may analyse the existence and uniqueness of solutions to the NTE; see [6] and [7]. Moreover, a similar approach has also been pioneered by Mokhtar-Kharroubi [27].

The probabilistic interpretation of the NTE was appreciated from the very first treatments of the NTE (see e.g. [8] and references therein as well as Bell [2]). Indeed, the physical description of nuclear fission, when governed by basic principles, allowing for additional randomness, is nothing more than a branching Markov process. Numerous derivations of the NTE from this perspective can be found in the literature to various degrees of rigour; see e.g. Bell [2], Mori et al. [28], Pazy and Rabinowitz, [31], Lewins [23] and Pázsit and Pál. [29].

A more modern treatment of the probabilistic representation through Feynman-Kac ex-

pectation semigroups and the connection to the theory of Markov diffusions is found in Dautray et al. [6]. A purely probabilistic treatment can be found in Lapeyre et al. [19]. See also the accompanying papers to this one [5, 15, 13].

We finish this section by noting that there is a body of literature that pertains to the numerical analysis of the NTE. Recent work in this field, including the notion of uncertainty quantification, can be found in e.g. [24, 16, 32]. See also references therein.

2.5 Multi-species (Backwards) Neutron Transport Equation

In the following discussion, rather than talk about typed particles, we prefer to say typed ‘emissions’ as the different types correspond to particles, electromagnetic rays (e.g. gamma rays) and isotopes (which are considered to be carriers for delayed emissions). Let us now introduce an advanced version of the NTE, which takes account of both non-transmutation emissions as well as transmutation emissions, in particular, allowing for the inclusion of all types of emissions, prompt neutrons, delayed neutrons, alpha, beta and gamma emissions etc. An important feature (and arguably a restriction) of our model is that only prompt neutrons can produce delayed emissions.

In order to keep track of the various emission types, we define the type space $I := \{1, \dots, m\}$ for some $m \in \mathbb{N}$, ordered such that

- type 1 emissions: prompt neutrons (neutrons released immediately after fission)
- types $2, \dots, \ell$ emissions: other prompt emissions (e.g. alpha, beta, gamma emissions)
- types $\ell + 1, \dots, m$ emissions: isotopes (holding types/precursors) that hold delayed emissions.

Finally, the set of admissible velocities for each of the types i can be embedded within a common space $V = \{v \in \mathbb{R}^3 : v_{\min} \leq |v| \leq v_{\max}\}$, with $0 < v_{\min} \leq v_{\max} < \infty$). In practice, neutrons travel with a minimum speed of $\frac{1}{40}$ MeV but from a mathematical perspective, a lower bound on the speed ensures irreducibility of the system, as we later discuss.

We now consider the flux, $\psi_t(i, r, v)$ of type i emissions through a given region $r \in D$ with velocity $v \in V$ at time $t \geq 0$. We are interested in the so called *multi-species neutron transport equation (MNTE)* which takes the form

$$\begin{aligned}
\frac{\partial}{\partial t} \psi_t(i, r, v) &= v \cdot \nabla \psi_t(i, r, v) - \sigma^i(r, v) \psi_t(i, r, v) \\
&+ \sigma_{\mathbf{s}}^i(r, v) \int_V \psi_t(i, r, v) \pi_{\mathbf{s}}^i(r, v, v') dv' \\
&+ \sigma_{\mathbf{f}}^i(r, v) \sum_{j=1}^{\ell} \int_V \psi_t(j, r, v) \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' \\
&+ \mathbf{1}_{(i=1)} \sigma_{\mathbf{f}}^1(r, v) \sum_{j=\ell+1}^m m^j(r, v) \psi_t(j, r, v), \tag{2.9}
\end{aligned}$$

for prompt emissions $i = 1, \dots, \ell$, whereas, in the case of delayed emissions, $i = \ell + 1, \dots, m$ satisfies

$$\frac{\partial}{\partial t} \psi_t(i, r, v) = -\lambda_i \psi_t(i, r, v) + \lambda_i \sum_{j=1}^{\ell} \int_V \psi_t(j, r, v) \pi_{\mathbf{f}}^{i,j}(r, v, v') dv', \quad (2.10)$$

which is of a simple form because it describes only how these emissions are held in a suspended state (no advection) before being converted back to prompt emissions. Similarly to before, we have the following interpretation:

- $\sigma_{\mathbf{s}}^i(r, v)$: the rate at which scattering occurs for a type i emission with incoming velocity v ,
- $\sigma_{\mathbf{f}}^i(r, v)$: the rate at which fission occurs for a type i emission with incoming velocity v ,
- $\sigma^i(r, v)$: the sum of the rates $\sigma_{\mathbf{f}}^i + \sigma_{\mathbf{s}}^i$ and is known as the total cross section for a type i emission,
- $\pi_{\mathbf{s}}^i(r, v, v') dv'$: the scattering yield at velocity v' from incoming velocity v for a type i emission, satisfying $\int_V \pi_{\mathbf{s}}^i(r, v, v') dv' = 1$,
- $\pi_{\mathbf{f}}^{i,j}(r, v, v') dv'$: the average type j yield at velocity v' from fission with incoming velocity v for a type i emission satisfying $\sum_{j=1}^{\ell} \int_V \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' < \infty$,
- $m^j(r, v)$: the average type j (unstable) isotope yield from a fission event due to a type 1 particle with incoming velocity v ,
- λ_i : the decay rate for a type i isotope.

The above quantities are also called cross sections and there are a number of assumptions about these quantities that will remain in force throughout the remainder of this text.

Assumption 1. All cross sections are non-negative, measurable and uniformly bounded from above. Moreover, all prompt emissions scatter and hence, without loss of generality, we also assume that for for each $i = 1, \dots, \ell$, the terms $\sigma_{\mathbf{s}}^i \pi_{\mathbf{s}}^i$ are uniformly bounded away from the origin on $D \times V \times V$. We need not assume that the cross sections $\sigma_{\mathbf{f}}^i \pi_{\mathbf{f}}^{i,j}$ are uniformly bounded away from the origin for $1 \leq i, j \leq \ell$, with the exception of $i = 1$, for which it only makes sense that $\sigma_{\mathbf{f}}^1 m^j$ is uniformly bounded away from 0 for each $j = \ell + 1, \dots, m$. Without loss of generality, we can assume that $0 < \lambda_{\ell+1} < \dots < \lambda_m$.

We also assume similar boundary conditions to the single-type case in the sense that emissions exiting the physical domain D are killed. That is to say

$$\begin{cases} \psi_0(i, r, v) = g(i, r, v) & \text{for } 1 \leq i \leq m, r \in D, v \in V, \\ \psi_t(i, r, v) = 0 & \text{for } 1 \leq i \leq \ell, r \in \partial D \text{ if } v \cdot \mathbf{n}_r > 0. \end{cases} \quad (2.11)$$

For the second condition, we will write $\psi_t|_{\partial D^+} = 0$, where $\partial D^+ = \{(i, r, v) \in \{1, \dots, \ell\} \times \partial D \times V : v \cdot \mathbf{n}_r > 0\}$

Classical literature suggests that one can integrate delayed neutrons into the setting of the NTE by adding an inhomogeneity corresponding to the integral of incoming delayed neutrons from time $-\infty$ to the present; see e.g. [8]. A vectorial representation such as the one above can be found, however, in the work of [27]. There, only one category of prompt emissions are considered with multiple species of delayed neutrons.

As before, let us define the multi-species backward transport, scattering and fission operators as they appear in MNTE (2.9) and (2.10), acting on $f \in \prod_{i=1}^m L_2(D \times V)$, so that, for $i = 1, \dots, m$,

$$\left\{ \begin{array}{l} \overleftarrow{\mathbf{T}}_i f(\cdot, r, v) := \mathbf{1}_{(1 \leq i \leq \ell)} v \cdot \nabla f(i, r, v) \\ \overleftarrow{\mathbf{S}}_i f(\cdot, r, v) := \mathbf{1}_{(1 \leq i \leq \ell)} \int_V [f(i, r, v') - f(i, r, v)] \sigma_{\mathbf{s}}^i(r, v) \pi_{\mathbf{s}}^i(r, v, v') dv' \\ \overleftarrow{\mathbf{F}}_i f(\cdot, r, v) := \mathbf{1}_{(1 \leq i \leq \ell)} \left(\sum_{j=1}^{\ell} \int_V f(j, r, v') \sigma_{\mathbf{f}}^j(r, v) \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' - \sigma_{\mathbf{f}}^i(r, v) f(i, r, v) \right) \\ \quad + \mathbf{1}_{(i=1)} \sum_{j=\ell+1}^m \sigma_{\mathbf{f}}^j(r, v) m^j(r, v) f(j, r, v) \\ \quad + \mathbf{1}_{(\ell+1 \leq i \leq m)} \left(\lambda_i \sum_{j=1}^{\ell} \int_V f(j, r, v') \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' - \lambda_i f(i, r, v) \right), \end{array} \right.$$

with zero action otherwise.

It is not often that MNTE is stated as above in (2.9) and (2.10) in existing literature; see e.g. [27] for presentation of the NTE in a similar vectorial format, which allows for only one category of prompt neutrons.

The requirement that all cross sections are uniformly bounded is by far not the weakest assumption we can make (see e.g. Chapter XXI of [7]).

The precise mathematical sense in which we must understand solutions to the coupled system (2.9) and (2.10) needs some discussion before we can proceed. To this end, we shall first introduce some notational conventions.

As alluded to above, we are interested in a vector space of functions, written as the column vector $g(\cdot) = (g(1, \cdot), \dots, g(m, \cdot))^T$, whose entries $g(i, \cdot) : D \times V \rightarrow [0, \infty)$, for each $i = 1, \dots, m$. More precisely we are interested in functions $f \in \prod_{j=1}^m L_2(D \times V)$, which is easily verified to be itself an L_2 space with inner product given by

$$\langle f, g \rangle = \sum_{i=1}^m (f, g)_i, \quad \text{where} \quad (f, g)_i = \int_{D \times V} f(i, r, v) g(i, r, v) dr dv. \quad (2.12)$$

Generally speaking, for a scalar quantity which is indexed by i , say $a(i)$, when written without

the index, we will understand it to be a column vector. Sometimes we will want to put $f \in \prod_{j=1}^m L_2(D \times V)$ on the diagonal of an $m \times m$ matrix, in which case we will write $\mathbf{diag}(f)$. For our transport, scattering and fission operators, we will understand $\overleftarrow{\mathbf{T}} = \mathbf{diag}(\overleftarrow{\mathbf{T}}_1, \dots, \overleftarrow{\mathbf{T}}_m)$; however, we will understand $\overleftarrow{\mathbf{F}}$ to be the matrix acting on vectors $f \in \prod_{j=1}^m L_2(D \times V)$, with i, j -th entry given by

$$\begin{aligned} \overleftarrow{\mathbf{F}}_{i,j} f(j, r, v) &:= \mathbf{1}_{(1 \leq i, j \leq \ell)} \left(\int_V f(j, r, v') \sigma_{\mathbf{f}}^i(r, v) \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' - \mathbf{1}_{(i=j)} \sigma_{\mathbf{f}}^i(r, v) f(i, r, v) \right) \\ &\quad + \mathbf{1}_{(i=1, \ell+1 \leq j \leq m)} \sigma_{\mathbf{f}}^i(r, v) m^j(r, v) f(j, r, v) \\ &\quad + \mathbf{1}_{(\ell+1 \leq i \leq m, 1 \leq j \leq \ell)} \left(\lambda_i \int_V f(j, r, v') \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' - \mathbf{1}_{(i=j)} \lambda_i f(i, r, v) \right). \end{aligned}$$

The operator $\overleftarrow{\mathbf{S}}$ can be handled similarly to $\overleftarrow{\mathbf{T}}$.

We are fundamentally interested in a *classical solution* to the so-called (initial-value) *abstract Cauchy problem* (ACP)

$$\begin{cases} \frac{\partial}{\partial t} u_t &= (\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}) u_t \\ u_0 &= g \end{cases} \quad (2.13)$$

where u_t is treated as a column vector belonging to the space $\prod_{j=1}^m L_2(D \times V)$, for $t \geq 0$. Specifically this means that $(u_t, t \geq 0)$ is continuously differentiable in this space. In other words, there exists a $\dot{\psi}_t \in \prod_{j=1}^m L_2(D \times V)$, which is time-continuous in $\prod_{j=1}^m L_2(D \times V)$ with respect to $\|\cdot\|_2$, such that $\lim_{h \rightarrow 0} h^{-1}(u_{t+h} - u_t) = \dot{\psi}_t$ for all $t \geq 0$.

The theory of c_0 -semigroups gives us a straightforward approach to describing the unique solution to (2.13). Recall that a c_0 -semigroup also goes by the name of a strongly continuous semigroup and, in the present context, this means a family of time-indexed operators, $(\mathbf{V}_t, t \geq 0)$, on $\prod_{j=1}^m L_2(D \times V)$ with the properties that

- (i) $\mathbf{V}_0 = \text{Id}$,
- (ii) $\mathbf{V}_{t+s}[g] = \mathbf{V}_t[\mathbf{V}_s[g]]$, for all $s, t \geq 0$, $g \in \prod_{j=1}^m L_2(D \times V)$ and
- (iii) for all $g \in \prod_{j=1}^m L_2(D \times V)$, $\lim_{h \rightarrow 0} \|\mathbf{V}_h[g] - g\|_2 = 0$.

To see how c_0 -semigroups relate to (2.13), let us define $\overleftarrow{\mathbf{A}} := \overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}$ and define $(\mathbf{V}_t[g], t \geq 0)$ the semigroup generated by $\overleftarrow{\mathbf{A}}$ via

$$\mathbf{V}_t[g] := \exp(t \overleftarrow{\mathbf{A}})g, \quad g \in \prod_{j=1}^m L_2(D \times V). \quad (2.14)$$

Note that

$$\text{Dom}(\overleftarrow{\mathbf{A}}) := \left\{ g \in \prod_{j=1}^m L_2(D \times V) : \lim_{h \rightarrow 0} h^{-1} \|\mathbf{V}_h[g] - g\|_2 \text{ exists} \right\}$$

is the domain of $\overleftarrow{\mathbb{A}}$ and standard theory (cf. [12]) tells us that $\mathbf{V}_t[g] \in \text{Dom}(\overleftarrow{\mathbb{A}})$ for all $t \geq 0$, and $g \in \text{Dom}(\overleftarrow{\mathbb{A}})$. Proposition II.6.2 of [12] now gives us the relevance to (2.13).

Theorem 2. (Proposition II.6.2, [12]) *Let $(\overleftarrow{\mathbb{A}}, \text{Dom}(\overleftarrow{\mathbb{A}}))$ be the generator of a c_0 -semigroup $(\mathbf{V}_t, t \geq 0)$. If $g \in \text{Dom}(\overleftarrow{\mathbb{A}})$, then $u_t := \mathbf{V}_t[g]$ is a representation of the unique classical solution of (2.13).*

The reader may well have wondered where the second boundary condition in (2.11) has gone in the above formulation. This is a matter of interpretation of $(\overleftarrow{\mathbb{T}}, \text{Dom}(\overleftarrow{\mathbb{T}}))$, and hence the generator $(\overleftarrow{\mathbb{A}}, \text{Dom}(\overleftarrow{\mathbb{A}}))$, as we now discuss.

We are interested in the advection semigroup with killing on the boundary of D ,

$$\mathbf{U}_t[g](i, r, v) = g(i, r + vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)}, \quad i = 1, \dots, \ell \text{ and } t \geq 0. \quad (2.15)$$

where

$$\kappa_{r,v}^D := \inf\{t > 0 : r + vt \notin D\}. \quad (2.16)$$

In essence, $(\mathbf{U}_s, s \geq 0)$ is the semigroup of the process which moves from a point of issue r in a straight line with velocity v and which is killed on hitting ∂D . To see why $\mathbf{U} := (\mathbf{U}_s, s \geq 0)$ has the semigroup property, note that

$$\kappa_{r+vs,v}^D = \inf\{t > 0 : r + v(t+s) \notin D\} = (\kappa_{r,v}^D - s) \vee 0,$$

so that $t < \kappa_{r+vs,v}^D$ if and only if $t + s < \kappa_{r,v}^D$. Hence for any $g \in \prod_{i=1}^m L_2(D \times V)$ satisfying the boundary conditions (2.11), we have from the definition (2.15), for $i = 1, \dots, \ell$, $r \in D$, $v \in V$,

$$\begin{aligned} \mathbf{U}_s[\mathbf{U}_t[g]](i, r, v) &= \mathbf{U}_t[g](i, r + vs, v) \mathbf{1}_{(s < \kappa_{r,v}^D)} \\ &= g(i, r + v(t+s), v) \mathbf{1}_{(t < \kappa_{r+vs,v}^D)} \mathbf{1}_{(s < \kappa_{r,v}^D)} \\ &= \mathbf{U}_{t+s}[g](i, r, v) \end{aligned}$$

Defining $\mathbf{U}_t[g](i, r, v) = g(i, r, v)$ for $i = \ell + 1, \dots, m$, it is a straightforward exercise, see e.g. Theorem 2 in Chapter XXI of [7], to show that \mathbf{U} is a c_0 -semigroup with generator $\overleftarrow{\mathbb{T}}$. Its domain satisfies

$$\begin{aligned} \text{Dom}(\overleftarrow{\mathbb{T}}) &= \prod_{i=1}^{\ell} \text{Dom}(\overleftarrow{\mathbb{T}}_i) \times \prod_{i=\ell+1}^m L_2(D \times V), \text{ where} \\ \text{Dom}(\overleftarrow{\mathbb{T}}_i) &= \left\{ g \in L_2(D \times V) : v \cdot \nabla g \in L_2(D \times V) \text{ and } g|_{\partial D^+} = 0 \right\}. \end{aligned} \quad (2.17)$$

Here, by $v \cdot \nabla g \in L_2(D \times V)$ we mean that $v \cdot \nabla g$ exists in the distributional sense and is integrable in the space $L_2(D \times V)$.

The domain of $\overleftarrow{\mathbb{A}}$ can be no larger than $\text{Dom}(\overleftarrow{\mathbb{T}})$. It turns out however that $\text{Dom}(\overleftarrow{\mathbb{A}}) = \text{Dom}(\overleftarrow{\mathbb{T}})$.

To see why, we need only consider that the linear operators of the form

$$K_i f(i, r, v) := \alpha^i(r, v) \sum_{j=1}^m \int_V f(j, r, v') \pi^{i,j}(r, v, v') dv,$$

are continuous mappings from $\prod_{i=1}^m L_2(D \times V)$ into itself, where α and $\pi^{i,j}$ are non-negative, measurable and uniformly bounded. The proof is a straightforward exercise which uses the Cauchy-Schwarz inequality; see for example Lemma XXI.1 of [7]. It follows that $\text{Dom}(\overleftarrow{\mathbf{S}})$ and $\text{Dom}(\overleftarrow{\mathbf{F}})$ are both equal to $\prod_{i=1}^m L_2(D \times V)$ and, hence, $\text{Dom}(\overleftarrow{\mathbf{A}})$ and $\text{Dom}(\overleftarrow{\mathbf{T}})$ agree.

Note there is no particular necessity to put solutions in an L_2 space. One might equally work with the space $\prod_{i=1}^m L_p(D \times V)$, for $p \in (1, \infty)$. As the reader might suspect, solutions of the backwards equation in an L_p space comes hand in hand with a similarly formulated solution to the forward equation in the conjugate space $\prod_{i=1}^m L_q(D \times V)$, where $q^{-1} + p^{-1} = 1$. See for example Chapter XXI of [7] or [27]. The reader will note the exclusion of the L_1 and L_∞ conjugacy. The reason for the exclusion boils down to the cumbersome nature of the advection operator $\overleftarrow{\mathbf{T}} = v \cdot \nabla$. Quite simply it is not possible to verify the strong continuity property of the advection semigroup

$$U_t[g](i, r, v) = g(i, r + vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)}, \quad t \geq 0. \quad (2.18)$$

where $\kappa_{r,v}^D := \inf\{t > 0 : r + vt \notin D\}$. Hence we cannot give a meaning to $v \cdot \nabla$ as a c_0 -semigroup on $L_\infty(D \times V)$. This is unfortunate as the latter is the more natural setting for probabilistic interpretation of solutions to the ACP. Having said that, the backwards scattering and fission operators, respectively $\overleftarrow{\mathbf{S}}$ and $\overleftarrow{\mathbf{F}}$, are well defined on all $\prod_{i=1}^m L_p(D \times V)$ spaces for $p \in [1, \infty]$.

One of our main results will be to establish the asymptotic (2.8) but now in the current setting. Recall that we have assumed that $D \subseteq \mathbb{R}^3$ is a smooth open pathwise connected bounded domain such that ∂D has zero Lebesgue measure.

Theorem 3. *Let D be convex. We assume the following irreducibility conditions. For each $i, j \in \{1, \dots, \ell\}$ assume that each of the cross sections $\sigma_{\mathbf{f}}^i(r, v) \pi_{\mathbf{f}}^{i,j}(r, v, v')$, $\sigma_{\mathbf{f}}^i(r, v) m^j(r, v)$ and $\sigma_{\mathbf{s}}^i(r, v) \pi_{\mathbf{s}}^i(r, v, v')$ are piece-wise continuous¹⁰ on $\bar{D} \times V \times V$ and there exists $k = k_{i,j} \in \{1, \dots, \ell\}$ such that*

$$\sigma_{\mathbf{f}}^i(r, v) \pi_{\mathbf{f}}^{i,k}(r, v, v') > 0 \text{ on } D \times V \times V \quad (2.19)$$

and

$$\sigma_{\mathbf{f}}^k(r, v) \pi_{\mathbf{f}}^{k,j}(r, v, v') > 0 \text{ on } D \times V \times V. \quad (2.20)$$

Then,

- (i) *the neutron transport operator $\overleftarrow{\mathbf{A}}$ has a simple and isolated eigenvalue $\lambda_c > -\lambda_{\ell+1}$, which is leading in the sense that $\lambda_c = \sup\{\text{Re}(\lambda) : \lambda \text{ is an eigenvalue of } \overleftarrow{\mathbf{A}}\}$ and which has*

¹⁰A function is piecewise continuous if its domain can be divided into an exhaustive finite partition (e.g. polytopes) such that there is continuity in each element of the partition. This is precisely how cross sections are stored in numerical libraries for modelling of nuclear reactor cores.

corresponding non-negative right and left eigenfunctions in $\prod_{i=1}^m L_2(D \times V)$, φ and $\tilde{\varphi}$ respectively and

(ii) there exists an $\varepsilon > 0$ such that, as $t \rightarrow \infty$,

$$\|e^{-\lambda_c t} \mathbf{V}_t[f] - \langle f, \tilde{\varphi} \rangle \varphi\|_2 = O(e^{-\varepsilon t}), \quad (2.21)$$

for all $f \in \prod_{i=1}^m L_2(D \times V)$, where $(\mathbf{V}_t, t \geq 0)$ is defined in (2.14). To give a precise value for ε , suppose we enumerate the real eigenvalues of $\overleftarrow{\mathbf{A}}$ in decreasing order by the set $\{\lambda^{(1)}, \dots, \lambda^{(n)}\}$ (noting from earlier that we have at least $\lambda^{(1)} = \lambda_c$). Then we have $\lambda^{(n)} > -\lambda_{\ell+1}$ and we can take any ε such that $\varepsilon < \lambda_c - (\lambda^{(2)} \vee (-\lambda_{\ell+1}))$ where we understand $\lambda^{(2)} = -\infty$ if $n = 1$.

Remark 1. It could be argued that the assumptions in the above theorem rule out the possibility that we may, for example, include alpha or beta emissions in the model for that particular conclusion. Whilst alpha and beta emissions may scatter, they are not energetic enough to cause fission. The irreducibility conditions (2.19) and (2.20) would thus fail. On the other hand, it is also known that when such particles are energetic enough, they can draw gamma radiation or positrons out of nuclei when passing in close proximity. If the latter are sufficiently energetic, then they can induce fission.

2.6 Multi-species neutron branching process

Heuristically speaking, (2.13) can be thought of as being closely related to the expectation semigroup of a Markov branching process, or *Multi-species nuclear branching process (MNBP)* as we shall call it, whose infinitesimal generator is $\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}$. Consider the system of typed emissions whose configurations in $D \times V$ at time $t \geq 0$ are given by $\{r_{i,j}(t), v_{i,j}(t) : i = 1, \dots, N_t^j\}$, where, for each $j = 1, \dots, m$, N_t^j is the number of type j emissions alive at time t . In order to describe the system as Markovian, we will represent it by the atomic measures

$$X_t(j, A) = \sum_{i=1}^{N_t^j} \delta_{(r_{i,j}(t), v_{i,j}(t))}(A), \quad j = 1, \dots, m,$$

where A is a Borel subset of $D \times V$ and δ is the Dirac measure defined on the same space. Then the system can be described via the m -tuple $X_t(\cdot) = (X_t(1, \cdot), \dots, X_t(m, \cdot))$, $t \geq 0$, which evolves as follows.

▷ A emission of type $i \in \{1, \dots, \ell\}$ with configuration (r, v) moves in a straight line with velocity v from the point r until one of the following events occur:

- The emission leaves the domain, at which point it is killed.
- Independently of all other emissions, a scattering event occurs when an emission comes in close proximity to an atomic nucleus and, accordingly, makes an instantaneous change

of velocity. For an emission in the system of type $i \in \{1, \dots, \ell\}$ with initial position and velocity (r, v) , if we write $T_{\mathbf{s}}^i$ for the random time until the next scattering occurs, then, independently of any other physical event that may affect the emission,

$$\Pr(T_{\mathbf{s}}^i > t) = \exp \left\{ - \int_0^t \sigma_{\mathbf{s}}^i(r + vs, v) ds \right\}. \quad (2.22)$$

- When scattering of an emission of type $i \in \{1, \dots, \ell\}$ occurs at space-velocity (r, v) , the new velocity is selected independently with probability $\pi_{\mathbf{s}}^i(r, v, v') dv'$.
- Independently of all other emissions, a fission event occurs when an emission smashes into an atomic nucleus. For an emission in the system with initial position and velocity (r, v) , we will write $T_{\mathbf{f}}^i$ for the random time that the next fission occurs. Then independently of any other physical event that may affect the emission,

$$\Pr(T_{\mathbf{f}}^i > t) = \exp \left\{ - \int_0^t \sigma_{\mathbf{f}}^i(r + vs, v) ds \right\}. \quad (2.23)$$

- When fission occurs, the smashing of the atomic nucleus releases a random number of other prompt emissions of type $i = 1, \dots, \ell$, say $N^{i,j} \geq 0$, which are ejected from the point of impact with randomly distributed, and possibly correlated, velocities, say $\{v_k^{i,j} : k = 1, \dots, N^{i,j}\}$. When fission occurs at location $r \in D$ from an emission with incoming velocity $v \in V$, the quantity $\pi_{\mathbf{f}}^{i,j}(r, v, v') dv'$ describes the average number of type j prompt emissions released from nuclear fission with outgoing velocity in the infinitesimal neighbourhood of v' . In particular

$$\int_A \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' = \mathbb{E} \left[\sum_{k=1}^{N^{i,j}} \mathbf{1}_{(v_k^{i,j} \in A)} \right], \quad A \in \mathcal{B}(V).$$

- Note, the possibility that $\Pr(N^{i,j} = 0) > 0$ is possible. If $i = j = 1$ then this is tantamount to neutron capture or further decomposition into subatomic particles which are not counted.
- Further, if the initial emission is a (type 1) neutron, a fission event (occurring at rate $\sigma_{\mathbf{f}}^1$) may result in the production of unstable isotopes (which later release delayed emissions). On this event, an average number, $m^j(r, v)$, of type $j \in \{\ell + 1, \dots, m\}$ isotopes will be produced from a collision at position r from a neutron with incoming velocity v . The isotopes will inherit the configuration of the incoming nucleus at the time of collision.

▷ An isotope of type $i \in \{\ell + 1, \dots, m\}$ with inherited physical configuration (r, v) stays in the same place for an exponentially distributed amount of time with rate λ_i . At this point, it produces a random number of type $j \in \{1, \dots, \ell\}$ prompt emissions, the average number of which, along with their corresponding velocities, are chosen according to $\pi_{\mathbf{f}}^{i,j}(r, v, v')$, in a

similar way to previously described. We note that although unstable isotopes stay in the same spatial position, we will still assign them a velocity as a ‘mark’.

In all cases, it is natural to make the following physical assumption which will remain in force throughout.

Assumption 4. Random emissions of any type are bounded in number by the non-random constant $n_{\max} \geq 1$. In particular this means that

$$\sup_{1 \leq i \leq m, 1 \leq j \leq \ell, r \in D, v \in V} \pi_{\mathbf{f}}^{i,j}(r, v, V) \leq n_{\max} \quad \text{and} \quad \sup_{r \in D, v \in V, 1 \leq j \leq \ell} m^j(r, v) \leq n_{\max}.$$

For non-negative and uniformly bounded $g : \prod_{i=1}^m (D \times V) \mapsto [0, \infty)$, that is $g \in \prod_{i=1}^m L_{\infty}^+(D \times V)$, define the *expectation semigroup*

$$\psi_t[g](i, r, v) := \mathbb{E}_{\delta_{(i,r,v)}}[\langle g, X_t \rangle], \quad (2.24)$$

where $\mathbb{P}_{\delta_{(i,r,v)}}$ is law of the process started from a single type i emission with configuration (r, v) with corresponding expectation operator $\mathbb{E}_{\delta_{(i,r,v)}}$.

As we have assumed that all cross sections are uniformly bounded, ignoring spatial trajectories of neutrons (in particular those that are killed by leaving the domain D), it is straightforward to compare the growth of $(\psi_t[g], t \geq 0)$ against that of a continuous-time Galton-Watson process with growth rate $\eta\{(m \times n_{\max}) - 1\}$, where $\eta = \sup_{1 \leq i \leq \ell, r \in D, v \in V} \sigma_{\mathbf{f}}^i(r, v) + \max_{\ell+1 \leq i \leq m} \lambda_i$.

The rate of growth $\eta\{(m \times n_{\max}) - 1\}$ simply assumes that each emission of type i gives rise to at most n_{\max} emissions of any other type and at a rate which is uniformly bounded by a uniform upper bound of all possible rates at which fission events occur. Note this rate takes account of the emission count introduced into the system at a fission event and the single emission removed from the system which caused the fission event.

It is also straightforward to stochastically upper bound the process $\langle 1, X_t \rangle$, $t \geq 0$, by the aforesaid continuous-time Galton Watson process on the same probability space. The latter process branches whenever X does, topping up the number of offspring always to n_{\max} , but also has additional independent branching events at rate $(\eta - \mathbf{1}_{(1 \leq i \leq \ell)} \sigma_{\mathbf{f}}^i(r, v) - \mathbf{1}_{(\ell+1 \leq i \leq m)} \lambda_i)$ always producing precisely n_{\max} offspring of each of the m possible emissions.

If we denote this Galton-Watson process by $(Z_t, t \geq 0)$, then we have both the stochastic bound $\langle 1, X_t \rangle \leq Z_t \leq Z_{t+s}$, for all $s, t \geq 0$ and the upper estimate

$$\sup_{1 \leq i \leq m, r \in D, v \in V} \psi_t[g](i, r, v) \leq \|g\|_{\infty} \exp(\eta((n_{\max} \times m) - 1)t), \quad t \geq 0. \quad (2.25)$$

If we put g in the smaller space $\prod_{i=1}^m C^+(D \times V)$, the space of non-negative, continuous and uniformly bounded vector functions on $D \times V$, then we also have by a dominated convergence argument, $\lim_{t \rightarrow 0} \psi_t[g] = g$ in the pointwise sense. Otherwise the latter convergence is not necessarily clear.

The name ‘expectation semigroup’ is earned thanks to the behaviour of $(\psi_t, t \geq 0)$ under an application of the Markov branching property. Indeed, associated to the MNBP are the probabilities \mathbb{P}_μ for atomic measures of the form

$$\mu = \left(\sum_{i=1}^{n_1} \delta_{(1,r_{i,1},v_{i,1})}, \dots, \sum_{i=1}^{n_m} \delta_{(m,r_{i,m},v_{i,m})} \right) =: (\mu_1, \dots, \mu_m). \quad (2.26)$$

The Markov branching property dictates that, for $g \in \prod_{i=1}^m L_2(D \times V)$ as before and $t \geq 0$,

$$\mathbb{E}_\mu[\langle g, X_t \rangle] = \sum_{j=1}^m \sum_{i=1}^{n_j} \mathbb{E}_{\delta_{(j,r_{i,j},v_{i,j})}}[\langle g, X_t \rangle] = \langle \mathbb{E}_{\delta_{(\cdot,\cdot)}}[\langle g, X_t \rangle], \mu \rangle$$

Here we are abusing our earlier notation in (2.12) and writing for finite atomic measures μ of the form (2.26),

$$\langle g, \mu \rangle = \sum_{i=1}^m (g, \mu)_i, \quad \text{where} \quad (g, \mu)_i = \int_{D \times V} g(i, r, v) \mu_i(dr, dv). \quad (2.27)$$

Hence, by conditioning on the configuration of the system at time $t \geq 0$, we have, for $s \geq 0$,

$$\psi_{t+s}[g](i, r, v) := \mathbb{E}_{\delta_{(i,r,v)}}[\mathbb{E}_{X_t}[\langle f, X_s \rangle]] = \mathbb{E}_{\delta_{(i,r,v)}}[\langle \psi_s[g], X_t \rangle] = \psi_t[\psi_s[g]](i, r, v). \quad (2.28)$$

The expectation semigroup property of $(\psi_t, t \geq 0)$ does not imply that it is necessarily a c_0 -semigroup on $\prod_{i=1}^m L_2(D \times V)$. Recalling our earlier discussion, if we were able to work with (2.13) in the setting of a c_0 -semigroup on $\prod_{i=1}^m L_\infty(D \times V)$, then we would be much closer to being able to match the expectation semigroup $(\psi_t, t \geq 0)$ to the solution $(u_t, t \geq 0)$. But even then, problems would occur with verifying strong continuity at the origin.

Nonetheless, classical literature supports the view that it is the physical processes, i.e. in this setting the MNBP, that provides a stochastic representation of the solution to the backward MNTE. The authors are not aware of a formal proof of this fact. We will nonetheless try to address this point shortly in Section 2.8. In the mean time, let us present an alternative ‘mild’ form of the MNTE (also called a *Duhamel solution* in the PDE literature) which the semigroup $(\psi_t, t \geq 0)$ more comfortably solves.

Lemma 1. *The expectation semigroup $(\psi_t[g], t \geq 0)$ is the unique solution in $\prod_{i=1}^m L_\infty^+(D \times V)$ to the mild MNTE*

$$u_t(i, r, v) = \mathbb{U}_t[g](i, r, v) + \int_0^t \mathbb{U}_s[(\tilde{\mathbb{S}} + \tilde{\mathbb{F}})u_{t-s}](i, r, v) ds, \quad (2.29)$$

for $t \geq 0$, $1 \leq i \leq m$, $r \in D$, $v \in V$ and $g \in \prod_{i=1}^m L_\infty^+(D \times V)$.

Before proceeding to the proof, let us remark that, in the statement of the theorem, we are not working with $(\mathbb{U}_t, t \geq 0)$ as a c_0 -semigroup on $\prod_{i=1}^m L_\infty(D \times V)$, but a pointwise shift operator. The reader will recall from the discussion preceding (2.18) that $(\mathbb{U}_t, t \geq 0)$ cannot be

defined as such for $\prod_{i=1}^m L_\infty(D \times V)$.

Proof of Lemma 1. First suppose we start with an emission of type i . By splitting the expectation in the definition of $\psi_t[g]$ at the first scattering or fission event, and remembering that the time $\kappa_{r,v}^D$ defined in (2.16) is deterministic, we have for $r \in D$ and $v \in V$,

$$\begin{aligned}
\psi_t[g](i, r, v) &= e^{-\int_0^t \sigma^i(r+vs, v) ds} g(i, r+vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)} \\
&\quad + \int_0^{t \wedge \kappa_{r,v}^D} \sigma^i(r+vs, v) e^{-\int_0^s \sigma^i(r+vu, v) du} \\
&\quad \left\{ \frac{\sigma_{\mathfrak{s}}^i(r+vs, v)}{\sigma^i(r+vs, v)} \int_V \psi_{t-s}[g](i, r+vs, v') \pi_{\mathfrak{s}}^i(r+vs, v, v') dv' \right. \\
&\quad \left. + \frac{\sigma_{\mathfrak{f}}^i(r+vs, v)}{\sigma^i(r+vs, v)} \left(\sum_{j=1}^m \int_V \psi_{t-s}[g](j, r+vs, v') \pi_{\mathfrak{f}}^{i,j}(r+vs, v, v') dv' \right. \right. \\
&\quad \left. \left. + \mathbf{1}_{(i=1)} \sum_{j=\ell+1}^m m^j(r+vs, v) \psi_{t-s}[g](j, r+vs, v) \right) \right\} ds \\
&= e^{-\int_0^t \sigma^i(r+vs, v) ds} g(i, r+vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)} \\
&\quad + \int_0^{t \wedge \kappa_{r,v}^D} e^{-\int_0^s \sigma^i(r+vu, v) du} (\overleftarrow{\mathfrak{S}}_i + \overleftarrow{\mathfrak{F}}_i + \sigma^i) \psi_{t-s}[g](i, r+vs, v) ds, \quad t \geq 0.
\end{aligned}$$

Now appealing to an analogue of Lemma 1.2, Chapter 4 in [11] (see also the Appendix of [18]), we can transfer the exponential integrals in each of the terms on the right-hand side above to a potential term in the integral as follows

$$\begin{aligned}
\psi_t[g](i, r, v) &= g(i, r+vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)} + \int_0^{t \wedge \kappa_{r,v}^D} (\overleftarrow{\mathfrak{S}}_i + \overleftarrow{\mathfrak{F}}_i + \sigma^i) \psi_{t-s}[g](i, r+vs, v) ds \\
&\quad - \int_0^t \sigma^i(r+vs, v) \psi_t[g](i, r, v) ds \\
&= g(i, r+vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)} + \int_0^{t \wedge \kappa_{r,v}^D} (\overleftarrow{\mathfrak{S}}_i + \overleftarrow{\mathfrak{F}}_i) \psi_{t-s}[g](i, r+vs, v) ds, \quad t \geq 0,
\end{aligned} \tag{2.30}$$

which agrees with (2.29), for $1 \leq i \leq \ell$.

Following a similar approach, for $\ell+1 \leq i \leq m$, $r \in D$, $v \in V$, we also get

$$\begin{aligned}
\psi_t[g](i, r, v) &= g(i, r, v) \mathbf{1}_{(t < \kappa_{r,v}^D)} - \lambda_i \int_0^{t \wedge \kappa_{r,v}^D} \psi_{t-s}[g](i, r, v) ds \\
&\quad + \int_0^{t \wedge \kappa_{r,v}^D} \lambda_i \left\{ \sum_{j=1}^{\ell} \int_V \psi_{t-s}[g](1, r, v') \pi_{\mathfrak{f}}^{i,j}(r, v, v') dv' \right\} ds \\
&= g(i, r, v) \mathbf{1}_{(t < \kappa_{r,v}^D)} + \int_0^{t \wedge \kappa_{r,v}^D} (\overleftarrow{\mathfrak{S}}_i + \overleftarrow{\mathfrak{F}}_i) \psi_{t-s}[g](i, r, v) ds, \quad t \geq 0,
\end{aligned} \tag{2.31}$$

noting in particular that, for $\ell + 1 \leq i \leq m$, $\overleftarrow{\mathbf{S}}_i \equiv 0$. Now putting (2.30) and (2.31) together we obtain (2.29).

For uniqueness, suppose that $(\psi_t^{(i)}, t \geq 0)$, $i = 1, 2$ are two bounded solutions to (2.29). Define $\chi_t[g] := |\psi_t^{(1)}[g] - \psi_t^{(2)}[g]|$ and note that, for $i = 1 \cdots, \ell$,

$$\begin{aligned}
\chi_t[g](i, r, v) &\leq \int_0^{t \wedge \kappa_{r,v}^D} |(\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})\psi_{t-s}^{(1)}[g](i, r + vs, v) - (\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})\psi_{t-s}^{(2)}[g](i, r + vs, v)| ds \\
&\leq \int_0^{t \wedge \kappa_{r,v}^D} (\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})|\psi_{t-s}^{(1)}[g](i, r + vs, v) - \psi_{t-s}^{(2)}[g](i, r + vs, v)| ds \\
&\leq \int_0^{t \wedge \kappa_{r,v}^D} (\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})\chi_{t-s}[g](i, r + vs, v) ds \\
&\leq C_1 \int_0^{t \wedge \kappa_{r,v}^D} \sum_{j=1}^m \int_V \chi_{t-s}[g](j, r + vs, v') dv' ds + C_2 \int_0^{t \wedge \kappa_{r,v}^D} \chi_{t-s}[g](i, r + vs, v) ds
\end{aligned} \tag{2.32}$$

for some constants $C_1, C_2 \in (0, \infty)$, where the final inequality follows on account of all cross sections being uniformly bounded.

Now define $\bar{\chi}_t[g] = \sup_{1 \leq i \leq m, r \in D, v \in V} \chi_t[g](i, r, v)$, $t \geq 0$. From (2.32) we have that

$$\bar{\chi}_t[g] \leq \left(C_1 \sum_{j=1}^m \mathbf{v} \circ \mathbf{1}(V) + C_2 \right) \int_0^t \bar{\chi}_{t-s}[g] ds. \tag{2.33}$$

Reversing the order of integration on the right-hand side above and then applying Grönwall's Lemma allows us to conclude that $\chi_t[g] \equiv 0$, which shows uniqueness. A similar argument holds for $i = \ell + 1, \dots, m$ \square

2.7 Multi-species neutron random walk and the Many-to-one Lemma

A second probabilistic perspective for analysing the MNTE is possible but seems rarely to have been discussed in existing literature, if at all. This consists of collapsing the sum of the operators $\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}$ to take the form $\overleftarrow{\mathbf{L}} + \mathbf{diag}(\beta)$ for an appropriate choice of β , where $\overleftarrow{\mathbf{L}}$ is the operator which is similar in structure to $\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}}$. In essence, this transformation, which we will describe more rigorously in a moment, heuristically postulates that the operator $\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}$ can be reinterpreted via a Feynman-Kac formula as the infinitesimal generator of a single emission which undergoes linear transport and scattering and which also accumulates potential β .

To describe this more precisely, we need to introduce the notion of a *multi-species neutron random walk* (MNRW). In the current setting this means a continuous-time typed random walk denoted by (J_t, R_t, Υ_t) , $t \geq 0$, on $\{1, \dots, m\} \times (D \times V)$ with additional cemetery state $\{\dagger\}$ when it exits the physical domain D or an emission otherwise disappears from the system. The MNRW is described by two fundamental quantities (which are functions of the current particle type, spatial

position and velocity). First, a scattering rate $\alpha^i(r, v)$, $i \in \{1, \dots, m\}$, $r \in D$, $v, v' \in V$, such that $\alpha^i(r, v) = \lambda_i$, for $i \in \{\ell + 1, \dots, m\}$. Second, a scattering probability kernel $\pi^{i,j}(r, v, v')$, $i, j \in \{1, \dots, m\}$, $r \in D$, $v, v' \in V$. In the spirit of the description of the MNBP, the MNRW is described as follows.

▷ When the MNRW is of type $i \in \{1, \dots, \ell\}$ with configuration (r, v) , it moves in a straight line with velocity v from the point r until one of the following events occur:

- When the MNRW position moves out of D or e.g. it decomposes into an emission type that is not counted, or is captured in a nucleus, it is instantaneously killed.
- A scattering event occurs and, accordingly, the MNRW keeps the same emission type but makes an instantaneous change of velocity. If we write $T_{\mathbf{s}}^i$ for the random time until the next scattering occurs, then,

$$\Pr(T_{\mathbf{s}}^i > t) = \exp \left\{ - \int_0^t \alpha^i(r + vs, v) ds \right\}. \quad (2.34)$$

- When scattering of an emission of type $i \in \{1, \dots, \ell\}$ occurs at space-velocity (r, v) , the new velocity and type is selected independently with probability $\pi^{i,j}(r, v, v') dv'$.

▷ Otherwise, if $\ell + 1 \leq i \leq m$, then the emission remains motionless, i.e. the random walk is dormant, holding its initial position r , but retaining the velocity v as a mark. After an independent and exponentially distributed random time with rate λ_i , the particle transfers its type to $j \in \{1, \dots, \ell\}$ and acquires a new velocity v' with probability density $\pi^{i,j}(r, v, v')$.

We can associate to the MNRW the infinitesimal generator

$$\begin{aligned} \tilde{\mathbf{L}}_i f(r, v) &:= \mathbf{1}_{(i \leq \ell)} v \cdot \nabla f(i, r, v) \mathbf{1}_{(r \in D)} \\ &+ \alpha^i(r, v) \sum_{j=1}^m \int_V [f(j, r, v') - f(i, r, v)] \pi^{i,j}(r, v, v') dv'. \end{aligned} \quad (2.35)$$

for $f \in \text{Dom}(\tilde{\mathbf{L}}) = \text{Dom}(\tilde{\mathbf{T}})$. We thus refer to the process as an $\tilde{\mathbf{L}}$ -MNRW.

With the notion of the MNRW in hand, let us consider the following algebraic manipulations. For $i \in \{1, \dots, \ell\}$, $j \in \{1, \dots, m\}$, $(r, v) \in D \times V$ and $v' \in V$, define

$$\begin{aligned} \alpha^i(r, v) &= \mathbf{1}_{(1 \leq i \leq \ell)} \sigma_{\mathbf{s}}^i(r, v) \\ &+ \mathbf{1}_{(1 \leq i \leq \ell)} \sigma_{\mathbf{f}}^i(r, v) \left(\sum_{j=1}^{\ell} \int_V \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' + \mathbf{1}_{(i=1)} \sum_{j=\ell+1}^m m^j(r, v) \right) \\ &+ \mathbf{1}_{(\ell+1 \leq i \leq m)} \lambda_i \sum_{j=1}^{\ell} \int_V \pi_{\mathbf{f}}^{i,j}(r, v, v') dv', \end{aligned} \quad (2.36)$$

$$\begin{aligned} \pi^{i,j}(r, v, v') = (\alpha^i(r, v))^{-1} & \left[\sigma_{\mathbf{s}}^i(r, v) \pi_{\mathbf{s}}^i(r, v, v') \mathbf{1}_{(1 \leq i=j \leq \ell)} \right. \\ & + \sigma_{\mathbf{f}}^i(r, v) \left(\pi_{\mathbf{f}}^{i,j}(r, v, v') \mathbf{1}_{(1 \leq i, j \leq \ell)} + m^j(r, v) \mathbf{1}_{(i=1, j > \ell)} \right) \\ & \left. + \lambda_i \pi_{\mathbf{f}}^{i,j}(r, v, v') \mathbf{1}_{(\ell+1 \leq i \leq m, j \leq \ell)} \right], \end{aligned} \quad (2.37)$$

$$\beta^i(r, v) = \alpha^i(r, v) - \mathbf{1}_{(1 \leq i \leq \ell)} \sigma_{\mathbf{s}}^i(r, v) - \mathbf{1}_{(\ell+1 \leq i \leq m)} \lambda_i - \mathbf{1}_{(1 \leq i \leq \ell)} \sigma_{\mathbf{f}}^i(r, v). \quad (2.38)$$

Note, in particular, that for each fixed $1 \leq i \leq m$, $r \in D$ and $v \in V$, $\pi^{i,j}(r, v, v')$ is a probability distribution on $\{1, \dots, m\} \times V$ in the sense that $\sum_{j=1}^m \int_V \pi^{i,j}(r, v, v') dv' = 1$. Note also that the assumption $\sum_{j=1}^{\ell} \int_V \pi_{\mathbf{f}}^{i,j}(r, v, v') dv' \geq 0$ ensures that $\beta^i \geq 0$, for $1 \leq i \leq m$.

With simple algebra, we may now identify

$$(\overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})f(r, v) = \overleftarrow{\mathbf{L}}f(r, v) + \text{diag}(\beta)f(r, v) \quad (2.39)$$

where $f \in \text{Dom}(\overleftarrow{\mathbf{A}})$ (for which it was remarked earlier that it is equal to $\text{Dom}(\overleftarrow{\mathbf{T}})$), and $\overleftarrow{\mathbf{L}}$ is given by (2.35).

Heuristically speaking, we have algebraically gathered all of the operators into the infinitesimal generator of an $\overleftarrow{\mathbf{L}}$ -MNRW and local potential β . This has the attraction of leading us the aforementioned single emission representation of the solution to the MNTE using a single-emission Feynman-Kac representation. Said another way, this means that one would expect that, in the appropriate sense, the solution to the NTE to be represented in the form

$$\phi_t[g](i, r, v) = \mathbf{E}_{(i,r,v)} \left[e^{\int_0^t \beta^{J_s}(R_s, \Upsilon_s) ds} g(J_t, R_t, \Upsilon_t) \mathbf{1}_{(t < \tau_D)} \right], \quad (2.40)$$

for $t \geq 0$, $1 \leq i \leq m$, $r \in D, v \in V$. Here $\mathbf{P}_{(i,r,v)}$ is the law of the $\overleftarrow{\mathbf{L}}$ -MNRW starting from a single emission with configuration (i, r, v) , and $\mathbf{E}_{(i,r,v)}$ is the corresponding expectation operator.

Appealing to the Markov property for (J, R, Υ) , it is not difficult to show that a semigroup property similar to (2.28) holds. That is to say, for $s, t \geq 0$, $1 \leq i \leq m$, $r \in D, v \in V$

$$\phi_{s+t}[g](i, r, v) = \phi_s[\phi_t[g]](i, r, v).$$

Similarly to the case of $(\psi_t[g], t \geq 0)$, if we put g in the smaller space $\prod_{i=1}^m C^+(D \times V)$ then we also have $\lim_{t \rightarrow 0} \phi_t[g] = g$ in the pointwise sense, but otherwise strong continuity at $t = 0$ is unclear. Note also that, since all cross sections are uniformly bounded, then so is β (in all of its variables) by a constant, say $\bar{\beta}$. Hence, for $g \in \prod_{i=1}^m L_\infty(D \times V)$, $\|\phi_t[g]\|_\infty \leq \|g\|_\infty \exp(\bar{\beta}t)$, $t \geq 0$. As with the case of $(\psi_t[g], t \geq 0)$, the notion that $(\phi_t[g], t \geq 0)$ solves (2.13) is not a straightforward claim. Nonetheless, as one might expect, these two expectation semigroups are equal and we can see this by relating back to (2.29).

Indeed, by conditioning the expectation in the definition of $\phi_t[g]$ on the first scattering event, and then appealing to the Lemma 1.2, Chapter 4 in [11] in a similar manner to what was

done in the proof of Lemma 1, one easily deduces the below result. In the the spatial branching process literature, this would be called a ‘many-to-one’ lemma.

Lemma 2. *For $g \in \prod_{i=1}^m L_\infty^+(D \times V)$, the two expectation semigroups $(\phi_t[g], t \geq 0)$ and $(\psi_t[g], t \geq 0)$ agree.*

2.8 Consolidating the ACP with the expectation semigroup

We want to understand how the $\prod_{i=1}^m L_2(D \times V)$ semigroup $(\mathbf{v}_t, t \geq 0)$ that represents the unique solution to the Abstract Cauchy Problem (2.13) relates to the expectation semigroups $(\psi_t, t \geq 0)$ and $(\phi_t, t \geq 0)$ that offer two different stochastic representations to the mild equation (2.29).

We start by noting that if $g \in \prod_{i=1}^m L_\infty^+(D \times V)$, then, on account of the fact that $\text{Vol}(\prod_{i=1}^m (D \times V)) = (\int_{D \times V} dr dv)^m < \infty$, we also have $g \in \prod_{i=1}^m L_2(D \times V)$. Since it is unclear whether $(\psi_t[g], t \geq 0)$ is well defined for all $g \in \prod_{i=1}^m L_2(D \times V)$, it makes sense to consider the comparison with $(\mathbf{v}_t[g], t \geq 0)$ (defined in (2.14)) for the more restrictive choice $g \in \prod_{i=1}^m L_\infty(D \times V)$. The natural setting in which to make the comparison is in the space $\prod_{i=1}^m L_2(D \times V)$ as, by (2.25), $\|\psi_t[g]\|_\infty < \infty$ and the latter implies $\|\psi_t[g]\|_2 < \infty$, again thanks to the fact that $\text{Vol}(\prod_{i=1}^m (D \times V)) < \infty$.

Theorem 5. *If $g \in \prod_{i=1}^m L_\infty^+(D \times V)$ then, for $t \geq 0$, $\mathbf{v}_t[g] = \psi_t[g]$ on $\prod_{i=1}^m L_2(D \times V)$, i.e. $\|\mathbf{v}_t[g] - \psi_t[g]\|_2 = 0$.*

Before moving to its proof, the reader should take care to note that this does not imply that $(\mathbf{v}_t, t \geq 0)$ and $(\psi_t, t \geq 0)$ agree as c_0 -semigroups on $\prod_{i=1}^m L_2(D \times V)$. In particular, the comparison between the two semigroup operators is only made on $\prod_{i=1}^m L_2(D \times V)$, and $(\psi_t, t \geq 0)$ was not (and in fact cannot be) shown to demonstrate the strong continuity property on $\prod_{i=1}^m L_2(D \times V)$.

Remark 2. If we consider Theorem 5 in light of Theorem 3, noting that $(\psi_t[g], t \geq 0)$ is a uniformly bounded sequence, it is tempting to want to say that the leading eigenfunction φ belongs to $\prod_{i=1}^m L_\infty(D \times V)$. This is not the case necessarily and remains to be proved. In the setting of a single type of emission, this will be demonstrated in the forthcoming paper [15].

Proof of Theorem 5. Consider the adjusted ACP with inhomogeneity given by

$$\begin{cases} \frac{\partial u_t}{\partial t} &= \bar{\mathbf{T}} u_t + (\bar{\mathbf{S}} + \bar{\mathbf{F}}) \mathbf{v}_t[g] \\ u_0 &= g \end{cases} \quad (2.41)$$

By taking the difference of two solutions and invoking the uniqueness of the ACP in $\prod_{i=1}^m L_2(D \times V)$ with initial data $g = 0$, we note that the solution to (2.41) is unique in $\prod_{i=1}^m L_2(D \times V)$. However, on the one hand, it is straightforward to verify that

$$u_t := e^{t\bar{\mathbf{T}}} g + \int_0^t e^{(t-s)\bar{\mathbf{T}}} (\bar{\mathbf{S}} + \bar{\mathbf{F}}) \mathbf{v}_s[g] ds, \quad t \geq 0,$$

solves (2.41). On the other hand, taking account of the fact that $(\mathbf{V}_t[g], t \geq 0)$ solves (2.13), it is also the case that taking

$$u_t := \mathbf{V}_t[g], \quad t \geq 0,$$

also gives a solution to (2.41). Uniqueness thus tells us that on $\prod_{i=1}^m L_2(D \times V)$,

$$\mathbf{V}_t[g] = \mathbf{U}_t[g] + \int_0^t e^{(t-s)\tilde{\mathbf{T}}} (\tilde{\mathbf{S}} + \tilde{\mathbf{F}}) \mathbf{V}_s[g] ds = \mathbf{U}_t[g] + \int_0^t \mathbf{U}_s[(\tilde{\mathbf{S}} + \tilde{\mathbf{F}}) \mathbf{V}_{t-s}[g]] ds, \quad t \geq 0,$$

where in the second equality we have reversed the direction of integration. In conclusion, whereas $(\psi_t[g], t \geq 0)$ solves (2.29) in the pointwise sense, $(\mathbf{V}_t[g], t \geq 0)$ solves it in the $\prod_{i=1}^m L_2(D \times V)$ sense.

On the other hand, we know that $(\psi_t[g], t \geq 0)$ is valued in $\prod_{i=1}^m L_2(D \times V)$, hence we can consider,

$$\|\psi_t[g] - \mathbf{V}_t[g]\|_2 = \left\| \int_0^t \mathbf{U}_s[(\tilde{\mathbf{S}} + \tilde{\mathbf{F}}) \{\psi_{t-s}[g] - \mathbf{V}_{t-s}[g]\}] ds \right\|_2, \quad t \geq 0.$$

To conclude the theorem, let us note that, for $T > 0$, and $w_t \in \prod_{i=1}^m L_2(D \times V)$, $t \leq T$, we have

$$\begin{aligned} \left\| \int_0^t w_s ds \right\|_2^2 &= \int_{D \times V} \left(t \int_0^t w_s(r, v) \frac{ds}{t} \right)^2 dr dv \\ &\leq \int_{D \times V} t^2 \left(\int_0^t w_s(r, v)^2 \frac{ds}{t} \right) dr dv \\ &\leq T \int_0^t \|w_s\|_2^2 ds, \quad t \leq T, \end{aligned} \tag{2.42}$$

where in the first inequality we have used Jensen's inequality and Cauchy-Schwarz in the second. Moreover, for $f \in \prod_{i=1}^m L_2(D \times V)$,

$$\begin{aligned} \|\mathbf{U}_s[f]\|_2^2 &= \sum_{i=1}^m \int_{D \times V} \mathbf{1}_{(s < \kappa_{r,v}^D)} f(i, r + vs, v)^2 dr dv \\ &\leq \sum_{i=1}^m \int_{D \times V} f(i, r', v)^2 dr' dv \\ &= \|f\|_2^2 \end{aligned} \tag{2.43}$$

where the inequality follows as a consequence that, for each v , the integral $\int_D \mathbf{1}_{(s < \kappa_{r,v}^D)} v(i, r + vs, v)^2 dr$ integrates over a subdomain of D . Also, we have for the operator $\tilde{\mathbf{S}}$ (and similarly for

$\check{\mathbf{F}})$, for $f \in \prod_{i=1}^m L_2(D \times V)$,

$$\begin{aligned}
\|(\check{\mathbf{S}} + \text{diag}(\sigma_s))f\|_2 &= \left(\sum_{i=1}^m \int_{D \times V} \left(\int_V f(i, r, v') \sigma_s(r, v) \pi_s^i(r, v, v') dv' \right)^2 dr dv \right)^{1/2} \\
&\leq C \left(\sum_{i=1}^m \int_{D \times V} \left(\int_V f(i, r, v') \times 1 dv' \right)^2 dr dv \right)^{1/2} \\
&\leq C \left(\sum_{i=1}^m \text{Vol}(V) \int_{D \times V} \int_V f(i, r, v')^2 dv' dr \right)^{1/2} \\
&\leq C \max_{1 \leq i \leq m} \text{Vol}(V) \|f\|_2,
\end{aligned} \tag{2.44}$$

where the constant C appears by upper estimating the uniformly bounded cross sections and in the second inequality we have used Cauchy-Schwarz.

It thus follows from (2.42), (2.43) and (2.44) that, for $t \leq T$, writing $\omega_t = \psi_t[g] - \mathbf{V}_t[g]$, $t \geq 0$,

$$\begin{aligned}
\|\omega_t\|_2^2 &= \left\| \int_0^t \mathbf{U}_s[(\check{\mathbf{S}} + \check{\mathbf{F}})\omega_{t-s}] ds \right\|_2^2 \\
&\leq T \int_0^t \|\mathbf{U}_s[(\check{\mathbf{S}} + \check{\mathbf{F}})\omega_{t-s}]\|_2^2 ds \\
&\leq T \int_0^t \|(\check{\mathbf{S}} + \check{\mathbf{F}})\omega_{t-s}\|_2^2 ds \\
&= T \int_0^t \|(\check{\mathbf{S}} + \check{\mathbf{F}} + \text{diag}(\sigma) - \text{diag}(\sigma))\omega_s\|_2^2 ds \\
&\leq T \int_0^t \left(\|(\check{\mathbf{S}} + \text{diag}(\sigma_s))\omega_s\|_2 + \|(\check{\mathbf{F}} + \text{diag}(\sigma_f))\omega_s\|_2 + \|\text{diag}(\sigma)\omega_s\|_2 \right)^2 ds \\
&\leq C' \int_0^t \|\omega_s\|_2^2 ds, \quad t \leq T,
\end{aligned} \tag{2.45}$$

where the constant C' comes from the fact that σ is uniformly bounded. The final inequality in (2.45) together with Grönwall's Lemma now tells us that $\|\omega_t\|_2 = 0$, for all $t \leq T$. Since T is chosen arbitrarily, it follows that $(\psi_t[g], t \geq 0)$ and $(\mathbf{V}_t[g], t \geq 0)$ are indistinguishable in $\prod_{i=1}^m L_2(D \times V)$. \square

The conclusion of this section is that it is not unreasonable to now understand the expectation semigroups $(\psi_t[g], t \geq 0)$ and $(\phi_t[g], t \geq 0)$ for non-negative, bounded and measurable g on $D \times V$ as the 'solution' to the MNTTE in place of $(\mathbf{V}_t[g], t \geq 0)$ for the same class of g . Indeed, the two agree both in $\prod_{i=1}^m L_2(D \times V)$ and hence $(dr \times dv)$ -Lebesgue almost everywhere.

The reader will also note that from the perspective of Monte Carlo simulation, the expectation semigroup $(\phi_t[g], t \geq 0)$ carries the potential to be exploited in a way that $(\psi_t[g], t \geq 0)$ cannot. More precisely, where branching trees are difficult to simulate and are not convenient for Monte Carlo computational parallelisation, random walks are. This simple idea is explored in greater detail in the accompanying paper to this one [5].

2.9 Asymptotic behaviour of the MNTE: Proof of Theorem 3

In this section we return to the fundamental notion that the solution to the MNTE in the form (2.13) is described by its leading asymptotics for large times. That is to say, we give the proof of Theorem 3. Our proof follows closely ideas found in Chapters 4 and 5 of [27].

Recall that the quantities α^i , $\pi^{i,j}$, β^i , $i, j = 1, \dots, m$ were defined in (2.36), (2.37) and (2.38) respectively. They were arranged into the operator $\overleftarrow{\mathbf{A}} = \overleftarrow{\mathbf{T}} + \overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}$, such that $\text{Dom}(\overleftarrow{\mathbf{A}}) = \text{Dom}(\overleftarrow{\mathbf{T}})$, described in (2.17).

For $j = 1, \dots, m$, let us introduce the operators $K_{i,j}$ on $L_2(D \times V)$ by

$$K_{i,j}f(r, v) = \alpha^i(r, v) \int_V f(r, v') \pi^{i,j}(r, v, v') dv'.$$

These are integral operators, which take the form

$$K_{i,j}f(r, v) = \int_V f(r, v') \mathbf{k}_{i,j}(r, v, v') dv'$$

on $D \times V \times V$, where

$$\mathbf{k}_{i,j}(r, v, v') = \sigma_{\mathbf{s}}^i \pi_{\mathbf{s}}^{i,j}(r, v, v') + \sigma_{\mathbf{f}}^i \pi_{\mathbf{f}}^{i,j}(r, v, v'). \quad (2.46)$$

A similar computation to (2.44) also shows that $K_{i,j}g \in L_2(D \times V)$ when $g \in L_2(D \times V)$. Then from (2.9) and (2.37), taking care to note the use of the indicators for the inclusion of terms for different indices, we can write, for $1 \leq i \leq \ell$, for $g \in \text{Dom}(\overleftarrow{\mathbf{A}})$,

$$\begin{aligned} \overleftarrow{\mathbf{A}}_i g(i, r, v) &= \overleftarrow{\mathbf{T}}_i g(i, r, v) - \sigma^i(r, v) g(i, r, v) \\ &+ \sum_{j=1}^{\ell} K_{i,j} g(j, r, v) + \mathbf{1}_{(i=1)} \sigma^1(r, v) \sum_{j=\ell+1}^m m^j(r, v) g(j, r, v) \end{aligned} \quad (2.47)$$

Moreover, for $\ell + 1 \leq i \leq m$,

$$\overleftarrow{\mathbf{A}}_i g(i, r, v) = -\lambda_i g(i, r, v) + \sum_{j=1}^{\ell} K_{i,j} g(j, r, v) \quad (2.48)$$

With this notation, write

$$\begin{aligned} \mathbf{T} &= \text{diag}(\overleftarrow{\mathbf{T}}_1 - \sigma^1, \dots, \overleftarrow{\mathbf{T}}_{\ell} - \sigma^{\ell}), \\ \mathbf{\Lambda} &= \text{diag}(\lambda_{\ell+1}, \dots, \lambda_m), \\ \mathbf{K}^{\circ} &= (K_{i,j}), \quad \text{for } i, j = 1, \dots, \ell, \\ \mathbf{M} &= (\mathbf{M}_{i,j}), \quad \text{where } \mathbf{M}_{i,j} = \sigma^1(r, v) m^j(r, v) \mathbf{1}_{(i=1)}, \text{ for } i = 1, \dots, \ell, j = \ell + 1, \dots, m, \\ \mathbf{K}_\circ &= (K_{i,j}), \quad \text{for } i = \ell + 1, \dots, m, j = 1, \dots, \ell. \end{aligned}$$

Then the abstract Cauchy problem (2.13) on $\prod_{j=1}^m L_2(D \times V)$ may now be written in matrix

form

$$\frac{\partial}{\partial t} u_t = \mathbf{A} u_t, \quad t \geq 0.$$

where $\mathbf{A} = \mathbf{T} + \mathbf{K}$ and

$$\mathbf{T} = \begin{bmatrix} \mathbb{T} & \mathbf{0} \\ \mathbf{0} & -\Lambda \end{bmatrix} \quad \text{and} \quad \mathbf{K} = \begin{bmatrix} \mathbb{K}^\circ & \mathbb{M} \\ \mathbb{K}_\circ & \mathbf{0} \end{bmatrix}.$$

The matrix \mathbf{T} is an operator on $\prod_{i=1}^m L_2(D \times V)$ with domain

$$\text{Dom}(\mathbf{T}) = \prod_{i=1}^{\ell} \text{Dom}(\overleftarrow{\mathbb{T}}_i) \times \prod_{i=\ell+1}^m L_2(D \times V)$$

which generates the strongly continuous (or c_0 -) semigroup $(\mathbb{U}_t^{\mathbf{T}}, t \geq 0)$ given by

$$\mathbb{U}_t^{\mathbf{T}}[g] = \begin{cases} e^{-\int_0^t \sigma^i(r+vs, v) ds} \mathbb{U}_t[g] & 1 \leq i \leq \ell \\ e^{-\lambda_i t} g & \ell + 1 \leq i \leq m, \end{cases} \quad (2.49)$$

for $g \in \prod_{i=1}^m L_2(D \times V)$.

In order to prove Theorem 3, we consider a different operator that is related to \mathbf{A} as follows. Consider the eigenvalue problem

$$\mathbf{A}\varphi = \lambda\varphi, \quad \lambda > -\lambda_{\ell+1}, \quad (2.50)$$

for $\varphi \in \prod_{i=1}^m L_2(D \times V)$. Write

$$\varphi^\circ(\cdot) = (\varphi(1, \cdot), \dots, \varphi(\ell, \cdot)) \quad \text{and} \quad \varphi_\circ(\cdot) = (\varphi(\ell+1, \cdot), \dots, \varphi(m, \cdot))$$

so that φ is the concatenation $(\varphi^\circ, \varphi_\circ)$. Separating this into prompt and delayed initial emissions, it can be written as

$$\begin{aligned} \mathbb{T}\varphi^\circ + \mathbb{K}^\circ\varphi^\circ + \mathbb{M}\varphi_\circ &= \lambda\varphi^\circ \\ \lambda\mathbb{I}_{m-\ell}\varphi_\circ &= -\Lambda\varphi_\circ + \mathbb{K}_\circ\varphi^\circ, \end{aligned}$$

where $\mathbb{I}_{m-\ell}$ is the $(m-\ell) \times (m-\ell)$ identity matrix. Rearranging the second equation, we get

$$\varphi_\circ = (\lambda\mathbb{I}_{m-\ell} + \Lambda)^{-1} \mathbb{K}_\circ\varphi^\circ \quad (2.51)$$

and substituting this into the first equation,

$$(\lambda\mathbb{I}_\ell - \mathbb{T})^{-1} \mathbb{K}^\circ(\lambda)\varphi^\circ = \varphi^\circ \quad \text{where} \quad \mathbb{K}^\circ(\lambda) = \mathbb{K}^\circ + \mathbb{M}(\lambda\mathbb{I}_{m-\ell} + \Lambda)^{-1} \mathbb{K}_\circ. \quad (2.52)$$

Our strategy is to show that there exists a λ_c such that $(\lambda_c\mathbb{I}_\ell - \mathbb{T})^{-1} \mathbb{K}^\circ(\lambda_c)$ has a leading eigenvalue 1, and that this is equivalent to λ_c being an eigenvalue of \mathbf{A} . The tool we shall use to do this is the Krein-Rutman Theorem, which we recall here for convenience in a format that is appropriate for our use; c.f. [7, p. 286].

Theorem 6 (Krein-Rutman Theorem). *Let X be a Banach space and suppose it contains a convex cone \mathcal{C} such that $\mathcal{C} - \mathcal{C} := \{h = f - g : f, g \in \mathcal{C}\}$ is dense in X . Suppose \mathcal{L} is a positive compact linear operator on X such that $r(\mathcal{L}) := \sup\{|\lambda| : \lambda \in \Sigma(\mathcal{L})\} > 0$, where $\Sigma(\mathcal{L})$ is the spectrum of the operator \mathcal{L} . Then $r(\mathcal{L})$ is an eigenvalue of \mathcal{L} with a corresponding positive eigenfunction.*

Our proof of Theorem 3 requires the following intermediary result below. Before stating it, the reader is reminded that $\lambda_{\ell+1}, \dots, \lambda_m$ are arranged so that $\lambda_{\ell+1}$ is the smallest. Thus, the condition $\lambda > -\lambda_{\ell+1}$ ensures that $\mathsf{K}^\circ(\lambda)$ is well defined. In particular, $(\lambda\mathsf{I}_{m-\ell} + \Lambda)$ is invertible.

Proposition 1. *Under the assumptions of Theorem 3, for each $\lambda > -\lambda_{\ell+1}$, $r((\lambda\mathsf{I}_\ell - \mathsf{T})^{-1}\mathsf{K}^\circ(\lambda))$ is the leading eigenvalue of $(\lambda\mathsf{I}_\ell - \mathsf{T})^{-1}\mathsf{K}^\circ(\lambda)$ with a corresponding positive eigenfunction φ_λ° .*

Proof. In relation to the Krein-Rutman theorem stated above, our Banach space is $X = \prod_{i=1}^m L_2(D \times V)$ and the corresponding cone is $\mathcal{C} = \prod_{i=1}^m L_2^+(D \times V)$. It is clear that this cone is convex, and since every L_2 function can be written as the difference of its positive and negative parts, \mathcal{C} satisfies the assumptions of the theorem. We now break the rest of the proof into a number of steps which are stated with a proof immediately afterwards.

Step 1. First we claim that $(\lambda\mathsf{I}_\ell - \mathsf{T})^{-1}\mathsf{K}^\circ(\lambda)$ is a compact operator.

Fix $1 \leq i, j \leq m$. By Fubini's Theorem we have that $r \mapsto \mathsf{K}_{i,j}f(r, v)$ is measurable for $g \in L_2(D \times V)$. The operators $\mathsf{K}_{i,j}$ are also integral operators and therefore are continuous on $L_2(V)$ and compact. The assumed piecewise continuity of the cross sections $\sigma_{\mathsf{s}}^i \pi_{\mathsf{s}}^i$ and $\sigma_{\mathsf{f}}^i \pi_{\mathsf{f}}^{i,j}$ and the boundedness of the domain V is sufficient to ensure that $r \mapsto \mathsf{K}_{i,j} \cdot (r, \cdot)$ is piecewise continuous under the operator norm from $L_2(D \times V)$ to $L_2(V)$ and hence $\{\mathsf{K}_{i,j} \cdot (r, \cdot) : r \in D\}$ forms a relatively compact set in the space of linear operators on $L_2(V)$. With these properties, the mapping $r \mapsto \mathsf{K}_{i,j} \cdot (r, \cdot)$, for $r \in D$, is said to be *regular*. One similarly (but more easily) shows that $r \mapsto \mathsf{M}_{i,j} \cdot (r, \cdot)$ is regular for $r \in D$ as operators on $L_2(V)$. By linearity, this implies that, for $1 \leq i, j \leq \ell$, the mapping $r \mapsto \mathsf{K}^\circ(\lambda)_{i,j}$ is also regular. We now conclude by [27, Theorem 4.1], which we state here for convenience.

Theorem. *Let $1 < p < \infty$ and let D be bounded and convex. We assume that $d\mu$ is such that the hyperplanes have zero $d\mu$ -measure and that the collision operator K is regular. Then $K(\lambda\mathsf{I} - \mathsf{T})^{-1}$ and $(\lambda\mathsf{I} - \mathsf{T})^{-1}K$ are compact in $L^p(D \times V, dx d\mu)$.*

In this case $d\mu$ is Lebesgue measure on V so the condition on the hyperplanes is satisfied. Hence, due to the above analysis $(\lambda\mathsf{I}_\ell - \mathsf{T})^{-1}\mathsf{K}^\circ(\lambda)$ is a compact operator.

Remark 3. It is precisely at the application of [27, Theorem 4.1] that we need the convexity of the domain D , as this is required within the aforesaid result.

Step 2. Next we show that $(\lambda\mathsf{I}_\ell - \mathsf{T})^{-1}\mathsf{K}^\circ(\lambda)$ is a positive irreducible operator.

Positivity is a straightforward consequence of the assumptions on the operators $\mathsf{K}_{i,j}$ and the form of the semigroup defined in (2.49). For irreducibility, it is enough to show that there exists an integer $n \geq 1$ such that $[(\lambda\mathsf{I}_\ell - \mathsf{T})^{-1}\mathsf{K}^\circ(\lambda)]^n f > 0$ for each $f \in \prod_{i=1}^\ell L_2^+(D \times V)$, cf.

[1]. Note, this latter property is the definition for $(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)$ to be positivity improving. To this end, note that due to (2.52), the entries of $\mathbf{K}^\circ(\lambda)(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)$ satisfy

$$[\mathbf{K}^\circ(\lambda)(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)]_{i,j} \geq [\mathbf{K}^\circ(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ]_{i,j} = \sum_{k=1}^{\ell} \mathbf{K}_{i,k}(\lambda\mathbf{I}_\ell - \mathbf{T})_{k,k}^{-1}\mathbf{K}_{k,j},$$

and that $\mathbf{K}_{i,k}(\lambda\mathbf{I}_\ell - \mathbf{T})_{k,k}^{-1}\mathbf{K}_{k,j}$ is an integral operator $L_2(D \times V) \rightarrow L_2(D \times V)$, $1 \leq i, j \leq \ell$, whose kernel is greater than or equal to

$$\int_0^\infty e^{-\lambda + \underline{\sigma}^k(\frac{r-r'}{t})t} \mathbf{k}_{i,k}\left(r', \frac{r-r'}{t}, v''\right) \mathbf{k}_{k,j}\left(r, v, \frac{r-r'}{t}\right) \frac{dt}{t^n}, \quad (2.53)$$

where $\underline{\sigma}^k(v) = \inf_{r \in D} \{\sigma^k(r, v)\}$. Note, in order to produce this estimate, the reader will note that $(\lambda\mathbf{I}_\ell - \mathbf{T})_{k,k}^{-1}$ is the resolvent of $(\mathbf{U}_t^{\mathbf{T}}, t \geq 0)$ in (2.49). If we choose the index k as in the assumptions (2.19) and (2.20) then the lower bound (2.53) ensures that $[\mathbf{K}^\circ(\lambda)(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)]_{i,j}$ is positivity improving. It follows that $[(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)]^2$ is also positivity improving and therefore $(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)$ is irreducible.

Step 3. We claim that there exists a non-negative eigenfunction $0 \neq \varphi_\lambda \in \prod_{i=1}^{\ell} L_2(D \times V)$ for the operator $(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)$ with eigenvalue that agrees with $r((\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda))$.

We use de Pagter's Theorem, cf. [9], which says that the spectral radius of a positive, compact, irreducible operator is strictly positive; that is to say $r((\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)) > 0$. In turn the Krein-Rutman Theorem 6 states that $r((\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda))$ is thus an eigenvalue for the operator $(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)$ with a corresponding non-negative eigenfunction φ_λ° . \square

Proof of Theorem 3. (i) In looking for a non-negative eigenfunction of $\overleftarrow{\mathbf{A}}$ with real eigenvalue, our earlier discussion tells us we must equivalently look for a solution to (2.50) and hence (2.52). This is equivalent to finding a real value λ_c such that $r((\lambda_c\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda_c)) = 1$. We again achieve this goal in steps.

Step 1. We want to show that

$$\lim_{\lambda \downarrow -\lambda_{\ell+1}} r((\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)) = \infty. \quad (2.54)$$

Recall that $(\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)$ is compact and irreducible so by [27, Theorem 5.13] we have the comparison of the spectral radii,

$$r((\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)) \geq r((\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\Delta[\mathbf{K}^\circ(\lambda)]), \quad (2.55)$$

where $\Delta[\mathbf{K}^\circ(\lambda)]$ is the matrix whose entries are given by $\Delta[\mathbf{K}^\circ(\lambda)] = \text{diag}(\mathbf{K}^\circ(\lambda)_{1,1}, \dots, \mathbf{K}^\circ(\lambda)_{\ell,\ell})$.

Suppose Δ is an $\ell \times \ell$ whose diagonal entries are given by operators Δ_i on $L_2(D \times V)$, for $i = 1, \dots, \ell$. If $\mu \in \Sigma(\Delta_1)$, the spectrum of Δ_1 , then $(\mu\mathbf{I}_\ell - \Delta)_{1,1}$ is not invertible, and so $\mu\mathbf{I}_\ell - \Delta$ is also not invertible. Hence $\mu \in \Sigma(\Delta)$, the spectrum of Δ , and so $\Sigma(\Delta_1) \subset \Sigma(\Delta)$.

Applying this argument to the diagonal matrix $(\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \Delta \mathbf{K}^\circ(\lambda)$, we have that

$$\Sigma([\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \Delta \mathbf{K}^\circ(\lambda)]_{1,1} \subset \Sigma((\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \Delta \mathbf{K}^\circ(\lambda)) \quad (2.56)$$

and so

$$r((\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \Delta \mathbf{K}^\circ(\lambda)) \geq r([\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \Delta \mathbf{K}^\circ(\lambda)]_{1,1} \geq r((\lambda - \overleftarrow{\mathbf{T}}_1 - \sigma^1)^{-1} \Delta [\mathbf{K}^\circ(\lambda)]_{1,1}). \quad (2.57)$$

where, in the final inequality, we have used (2.56).

Next recall that $(\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \mathbf{K}^\circ(\lambda) \varphi^\circ = \varphi^\circ$ where $\mathbf{K}^\circ(\lambda) = \mathbf{K}^\circ + \mathbf{M}(\lambda \mathbf{I}_{m-\ell} + \Lambda)^{-1} \mathbf{K}_\circ$. Similar reasoning to the proofs of previous steps shows us that $(\lambda - \overleftarrow{\mathbf{T}}_1 - \sigma^1)^{-1} \Delta [\mathbf{K}^\circ(\lambda)]_{1,1}$ and $(\lambda - \overleftarrow{\mathbf{T}}_1 - \sigma^1)^{-1} \sigma_{\mathbf{f}}^1 m^{\ell+1} (\mathbf{K}_\circ)_{1,\ell+1}$ are both compact and irreducible operators, so that

$$r((\lambda - \overleftarrow{\mathbf{T}}_1 - \sigma^1)^{-1} \Delta [\mathbf{K}^\circ(\lambda)]_{1,1}) \geq \frac{r((\lambda - \overleftarrow{\mathbf{T}}_1 - \sigma^1)^{-1} \sigma_{\mathbf{f}}^1 m^{\ell+1} (\mathbf{K}_\circ)_{1,\ell+1})}{\lambda + \lambda_{\ell+1}} > 0, \quad (2.58)$$

where the first inequality follows from [27, Theorem 5.13] and the second follows from [27, Theorem 5.7]. Combining (2.55), (2.57) and (2.58), we have

$$r((\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \mathbf{K}^\circ(\lambda)) \geq \frac{r((\lambda - \overleftarrow{\mathbf{T}}_1 - \sigma^1)^{-1} \sigma_{\mathbf{f}}^1 m^{\ell+1} (\mathbf{K}_\circ)_{1,\ell+1})}{\lambda + \lambda_{\ell+1}} > 0,$$

with the latter term tending to ∞ as $\lambda \rightarrow -\lambda_{\ell+1}$.

Step 2. Next we need to show that

$$\lim_{\lambda \rightarrow \infty} r((\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \mathbf{K}^\circ(\lambda)) < 1.$$

The spectral radius $r((\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \mathbf{K}^\circ(\lambda))$ as is $\mathbf{K}^\circ(\lambda)$. Using the standard operator norm $\|\cdot\|_2$ on $\prod_{i=1}^\ell L_2(D \times V)$,

$$\|\mathbf{K}^\circ(\lambda)g\|_2 = \|\mathbf{M}(\text{diag}((\lambda + \lambda_{\ell+1})^{-1}, \dots, (\lambda + \lambda_m)^{-1})) \mathbf{K}_\circ g\|_2$$

and, hence, by inspection, $\mathbf{K}^\circ(\lambda)$ is decreasing with λ and tends to \mathbf{K}° as $\lambda \rightarrow \infty$. Note, moreover, that for all $g \in \prod_{i=1}^\ell L_2(D \times V)$,

$$(\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \mathbf{K}^\circ(\lambda)g = \int_0^\infty e^{-\lambda t} \langle f, \mathbf{U}_t^T [\mathbf{K}^\circ(\lambda)g] \rangle dt,$$

showing similarly that $(\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \mathbf{K}^\circ(\lambda)$ is decreasing in λ . Due to [27, Lemma 8.1] (note that it is not difficult to see from the proof of that lemma that that the order of the operators there can be reversed), we have

$$\lim_{\lambda \rightarrow \infty} r((\lambda \mathbf{I}_\ell - \mathbf{T})^{-1} \mathbf{K}^\circ(\lambda)) < 1.$$

Step 3. In this penultimate step, we show that we have found a non-negative function of

$\overleftarrow{\mathbf{A}}$, with eigenvalue λ_c .

We have the existence of a $\lambda_c > -\lambda_{\ell+1}$ such that $r((\lambda\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(\lambda)) = 1$. That is to say, thanks to Proposition 1, we have found $\varphi^\circ = \varphi_{\lambda_c}^\circ$ which solves (2.52), which in turn, thanks to (2.51) gives us that $\varphi_\circ = (\lambda\mathbf{I}_{m-\ell} + \Lambda)^{-1}\mathbf{K}_\circ\varphi_{\lambda_c}^\circ$ so that with the concatenation

$$\varphi = (\varphi_{\lambda_c}^\circ, (\lambda_c\mathbf{I}_{m-\ell} + \Lambda)^{-1}\mathbf{K}_\circ\varphi_{\lambda_c}^\circ) \geq 0$$

we have the eigensolution

$$\mathbf{A}\varphi = \lambda_c\varphi.$$

which is equivalent to $\overleftarrow{\mathbf{A}}\varphi = \lambda_c\varphi$.

Step 4. For the final step we need to show that λ_c is the leading real eigenvalue of $\overleftarrow{\mathbf{A}}$, i.e.

$$\lambda_c = s(\mathbf{A}) := \sup\{\operatorname{Re}(\lambda) : \lambda \in \sigma(\mathbf{A})\},$$

where $\sigma(\mathbf{A})$ is the spectrum of the operator \mathbf{A} or equivalently of $\overleftarrow{\mathbf{A}}$. Moreover we need to show that it is simple and isolated.

We first note that since we have shown that $\lambda_c \in \sigma(\mathbf{A})$, in particular that the spectrum is non-empty, it follows from [27, Theorem 5.2] that $s(\mathbf{A}) \in \sigma(\mathbf{A})$. Now suppose that $\lambda_c \neq s(\mathbf{A})$ so that, in particular, $\lambda_c < s(\mathbf{A})$. Then, thanks again to [27, Lemma 8.1], $r((s(\mathbf{A})\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}^\circ(s(\mathbf{A}))) < 1$ and so 1 is not an eigenvalue of $(s(\mathbf{A})\mathbf{I}_\ell - \mathbf{T})^{-1}\mathbf{K}(s(\mathbf{A}))$. Said another way, this means that $s(\mathbf{A})$ is not an eigenvalue of \mathbf{A} (and hence of $\overleftarrow{\mathbf{A}}$), leading to a contradiction. Algebraic and geometric simplicity of λ_c follows from [7, Remark 12] and [7, Theorem 7(iii)], respectively. \square

Before turning to the proof of Theorem 3 (ii), we must state another intermediary result which is translated from a general setting of Banach operators to our current situation; cf. [27, Theorem 4.1] and [1, p. 359, Theorem 22].

Proposition 2. *Under the assumptions of Theorem 3*

$$\sigma(\mathbf{A}) \cap \{\operatorname{Re}(\lambda) : \lambda > s(\mathbf{T})\}$$

consists of isolated eigenvalues with finite multiplicities, where $s(\mathbf{T}) := \sup\{\operatorname{Re}(\lambda) : \lambda \in \sigma(\mathbf{T})\}$.

Note the Theorem from which the above proposition is derived in [1, p. 359, Theorem 22] requires as a sufficient condition that $(\lambda\mathbf{I} - \mathbf{T})^{-1}\mathbf{K}$ is compact, where \mathbf{I} is an $m \times m$ identity matrix. This fact easily follows from the conclusion in Step 1 of the proof of Proposition 1.

Finally we can complete the proof of Theorem 3

Proof of Theorem 3. (ii) It is also easy from the structure of \mathbf{T} that $-\lambda_{\ell+1}, \dots, -\lambda_m$, belong to its spectrum. Moreover, for all $i = 1, \dots, \ell$, $s(\overleftarrow{\mathbf{T}}_i - \sigma^i) = -\infty$. Since $-\lambda_{\ell+1}$ is the largest of these eigenvalues, and $\lambda_c > -\lambda_{\ell+1}$ (from part (i) of Theorem 3), Proposition 2 tells us that

$\sigma(\mathbf{A}) \cap \{\lambda : \operatorname{Re}(\lambda) > -\lambda_{\ell+1}\}$ contains at least one isolated eigenvalue with finite (algebraic) multiplicity (i.e. the lead eigenvalue λ_c).

Suppose we enumerate the eigenvalues in $\sigma(\mathbf{A}) \cap \{\lambda : \operatorname{Re}(\lambda) > -\lambda_{\ell+1}\}$ in decreasing order by the set $\{\lambda^{(1)}, \dots, \lambda^{(n)}\}$ (noting from earlier that we have at least $\lambda^{(1)} = \lambda_c$ and $\lambda^{(n)} > -\lambda_{\ell+1}$). Then, from [7, p. 265], for $g \in \operatorname{Dom}(\overleftarrow{\mathbf{A}})$, we have

$$V_t[g] = \sum_{k=1}^n e^{\lambda^{(k)}t} \left(\sum_{m=0}^{\operatorname{order}(\lambda^{(k)})-1} t^m \Pi_k^m[g] \right) + O(e^{-\lambda_{\ell+1}t}),$$

as $t \rightarrow \infty$, where Π_k are projectors in $\operatorname{Dom}(\overleftarrow{\mathbf{A}})$.

We are really only interested in the projection onto the eigenfunction that we know exists in the real part of the spectrum. The projector Π_1 can be written in the form

$$\Pi_1[g] = \langle g, \tilde{\varphi} \rangle \varphi, \quad g \in \prod_{i=1}^m L_2(D \times V),$$

where $\tilde{\varphi}$ is the left-eigenfunction with eigenvalue λ_c , which is guaranteed to exist by examining the preceding arguments for $\overleftarrow{\mathbf{A}}$ and re-applying them for $\overrightarrow{\mathbf{A}} := \overrightarrow{\mathbf{T}} + \overrightarrow{\mathbf{S}} + \overrightarrow{\mathbf{F}}$, the adjoint operator of $\overleftarrow{\mathbf{A}}$. Hence, we have the following leading order expansion,

$$V_t[f] = e^{\lambda_c t} (f, \tilde{\varphi}) \varphi + O(e^{[\lambda^{(2)} \vee (-\lambda_{\ell+1})]t}).$$

Note that since, according to Proposition 2, λ_c is isolated, there exists a $\varepsilon > 0$ such that $\lambda^{(2)} \vee (-\lambda_{\ell+1}) < \lambda_c - \varepsilon$, where we understand $\lambda^{(2)} = -\infty$ if $n = 1$. The statement of part (ii) of Theorem 3 now follows. \square

Acknowledgements

We are indebted to Paul Smith and Geoff Dobson from the ANSWERS modelling group at *Wood* for the extensive discussions as well as hosting at their offices in Dorchester. We would also like to thank Minmin Wang, Ivan Graham, Matt Parkinson and Denis Villemonais for useful discussions. Finally we would like to thank an enthusiastic referee for their comments and support of this article which is part review, part new results.

Bibliography

- [1] J. Banasiak and M. Mokhtar-Kharroubi, editors. *Evolutionary equations with applications in natural sciences*, volume 2126 of *Lecture Notes in Mathematics*. Springer, Cham, 2015. Lectures from the 2013 CIMPA-UNESCO-South Africa School held in Muizenberg, July 22–August 2, 2013.
- [2] G. I. Bell. On the stochastic theory of neutron transport. *Nuc. Sci. & Eng.*, 21:390–401, 1965.
- [3] A. Bensoussan, J.-L. Lions, and G. C. Papanicolaou. Asymptotics for branching transport processes. In *Computing methods in applied sciences and engineering (Proc. Third Internat. Sympos., Versailles, 1977), I*, volume 704 of *Lecture Notes in Math.*, pages 317–329. Springer, Berlin, 1979.
- [4] A. Bensoussan, J.-L. Lions, and G. C. Papanicolaou. Boundary layers and homogenization of transport processes. *Publ. Res. Inst. Math. Sci.*, 15(1):53–157, 1979.
- [5] A. M. G. Cox, S.C. Harris, A.E. Kyprianou, and M. Wang. Monte Carlo methods for the neutron transport equation. *Working document*, 2018.
- [6] 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 à l'énergie atomique. Eyrolles, Paris, 1989.
- [7] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 6*. Springer-Verlag, Berlin, 1993. Evolution problems. II, With the collaboration of Claude Bardos, Michel Cessenat, Alain Kavenoky, Patrick Lascaux, Bertrand Mercier, Olivier Pironneau, Bruno Scheurer and Rémi Sentis, Translated from the French by Alan Craig.
- [8] B. Davison and J. B. Sykes. *Neutron transport theory*. Oxford, at the Clarendon Press, 1957.
- [9] B. de Pagter. Irreducible compact operators. *Math. Z.*, 192(1):149–153, 1986.
- [10] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4):1671–1691, 1995.

- [11] E. B. Dynkin. *Diffusions, superdiffusions and partial differential equations*, volume 50 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2002.
- [12] K.-J. Engel and R. Nagel. *A short course on operator semigroups*. Universitext. Springer, New York, 2006.
- [13] S.C. Harris, E. Horton, and A.E. Kyprianou. Stochastic analysis of the neutron transport equation ii: Almost sure growth. *Preprint*, 2019.
- [14] T. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)].
- [15] E. Horton, A.E. Kyprianou, and D. Villemonais. Stochastic methods for the neutron transport equation i: Linear semigroup asymptotics. *Preprint*, 2019.
- [16] M.J. Parkinson I.G. Graham and R. Scheichl. Modern Monte Carlo variants for uncertainty quantification in neutron transport. *Festschrift for the 80th Birthday of Ian Sloan, eds. J. Dick, F.Y. Kuo, and H. Wozniakowski*, 2018.
- [17] K. Jörgens. An asymptotic expansion in the theory of neutron transport. *Comm. Pure Appl. Math.*, 11:219–242, 1958.
- [18] A. E. Kyprianou and S. Palau. Extinction properties of multi-type continuous-state branching processes. *Stochastic Process. Appl.*, 128(10):3466–3489, 2018.
- [19] B. Lapeyre, É. Pardoux, and R. Sentis. *Introduction to Monte-Carlo methods for transport and diffusion equations*, volume 6 of *Oxford Texts in Applied and Engineering Mathematics*. Oxford University Press, Oxford, 2003. Translated from the 1998 French original by Alan Craig and Fionn Craig.
- [20] J. Lehner. The spectrum of the neutron transport operator for the infinite slab. *J. Math. Mech.*, 11:173–181, 1962.
- [21] J. Lehner and G. M. Wing. On the spectrum of an unsymmetric operator arising in the transport theory of neutrons. *Comm. Pure Appl. Math.*, 8:217–234, 1955.
- [22] J. Lehner and G. M. Wing. Solution of the linearized Boltzmann transport equation for the slab geometry. *Duke Math. J.*, 23:125–142, 1956.
- [23] J. Lewins. Linear stochastic neutron transport theory. *Proc. Roy. Soc. London Ser. A*, 362(1711):537–558, 1978.
- [24] S. Maire and D. Talay. On a Monte Carlo method for neutron transport criticality computations. *IMA J. Numer. Anal.*, 26(4):657–685, 2006.

- [25] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.*, 25(3):487–517, 1993.
- [26] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.
- [27] M. Mokhtar-Kharroubi. *Mathematical topics in neutron transport theory*, volume 46 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co., Inc., River Edge, NJ, 1997. New aspects, With a chapter by M. Choulli and P. Stefanov.
- [28] T. Mori, S. Watanabe, and T. Yamada. On neutron branching processes. *Publ. Res. Inst. Math. Sci.*, 7:153–179, 1971/72.
- [29] I. Pázsit and L. Pál. *Neutron Fluctuations: A Treatise on the Physics of Branching Processes*. Elsevier, 2008.
- [30] A. Pazy and P. H. Rabinowitz. A nonlinear integral equation with applications to neutron transport theory. *Arch. Rational Mech. Anal.*, 32:226–246, 1969.
- [31] A. Pazy and P. H. Rabinowitz. On a branching process in neutron transport theory. *Arch. Rational Mech. Anal.*, 51:153–164, 1973.
- [32] F. Scheben and I. G. Graham. Iterative methods for neutron transport eigenvalue problems. *SIAM J. Sci. Comput.*, 33(5):2785–2804, 2011.

Concluding remarks

In this chapter we have addressed both the classical PDE perspective and the probabilistic branching perspective of the neutron transport equation, and consolidated these approaches in an appropriate sense. This will allow us to work with both approaches in later chapters.

We then considered a multi-type neutron transport equation that takes into account the full range of emissions associated with fission processes. Exploiting methods from spectral theory, we were then able to prove the existence of the leading eigenelements of this multi-type model, in the classical sense.

In the next chapter, we will focus on the branching process associated with the neutron transport equation, in order to answer the same questions but with probabilistic techniques.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:		
Stochastic Methods for the Neutron Transport Equation I: Linear Semigroup asymptotics		
Publication status (tick one)		
Draft manuscript	<input type="checkbox"/>	Submitted <input type="checkbox"/>
In review	<input checked="" type="checkbox"/>	Accepted <input type="checkbox"/>
Published	<input type="checkbox"/>	
Publication details (reference)	Under minor review for Annals of Applied Probability. AAP1901-001 <u>arXiv:1810.01779</u>	
Copyright status (tick the appropriate statement)		
I hold the copyright for this material	<input type="checkbox"/>	Copyright is retained by the publisher, but I have been given permission to replicate the material here <input checked="" type="checkbox"/>
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	<p>*All research has been conducted in equal partnership with collaborators and supervisors. It is unwise to try and measure percentages of intellectual contribution as research scholarship lies as much in the escalation of ideas through mathematical discourse as it does with the seed of ideas themselves. That said, the mathematical content of this thesis is, as an entire piece of work, inextricably associated to its author through intellectual ownership.</p> <p>Formulation of ideas: The candidate played an integral and fully collaborative role in the formulation of ideas.</p> <p>Design of methodology: The candidate played an integral and fully collaborative role in the design of methodology.</p> <p>Experimental work: N/A</p> <p>Presentation of data in journal format: N/A</p> <p>*The wording in this box follows the advice and approval of my supervisor, Professor Kyprianou.</p>	
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.	
Signed	Emma Horton	19/11/2019
		Date

Chapter 3

Linear semigroup asymptotics

Emma Horton¹, Andreas E. Kyprianou², Denis Villemonais³.

Remark 4. *This article has been submitted using the name *Stochastic Methods for the Neutron Transport Equation I: Linear Semigroup Asymptotics*.*

Abstract

The Neutron Transport Equation (NTE) describes the flux of neutrons through an inhomogeneous fissile medium. In this paper, we reconnect the NTE to the physical model of the spatial Markov branching process which describes the process of nuclear fission, transport, scattering, and absorption. By reformulating the NTE in its mild form and identifying its solution as an expectation semigroup, we use modern techniques to develop a Perron-Fröbenius (PF) type decomposition, showing that growth is dominated by a leading eigenfunction and its associated left and right eigenfunctions. In the spirit of results for spatial branching and fragmentation processes, we use our PF decomposition to show the existence of an intrinsic martingale and associated spine decomposition. Moreover, we show how criticality in the PF decomposition dictates the convergence of the intrinsic martingale. The mathematical difficulties in this context come about through unusual piecewise deterministic motion of particles coupled with an infinite type-space which is taken as neutron velocity. The fundamental nature of our PF decomposition also plays out in accompanying work [20, 9].

3.1 Introduction

The neutron transport equation (NTE) describes the flux of neutrons across a planar cross-section in an inhomogeneous fissile medium (measured in number of neutrons per cm^2 per

¹Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: elh48@bath.ac.uk

²Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: a.kyprianou@bath.ac.uk

³Institut Élie Cartan de Lorraine, Bureau 123, Université de Lorraine, 54506, Vandoeuvre-lès-Nancy Cedex, France. Email: denis.villemonais@univ-lorraine.fr

second). Neutron flux is described as a function of time, t , Euclidian location, r , direction, Ω and neutron energy E . It is not uncommon in the physics literature to assume that velocity is a function of both direction and energy, thereby reducing the number of variables by one. This allows us to describe the dependency of flux more simply in terms of time and, what we call, the *configuration variables* $(r, v) \in D \times V$ where $D \subseteq \mathbb{R}^3$ is a non-empty, open, smooth, bounded and convex domain such that ∂D has zero Lebesgue measure, and V is the velocity space, which we take to be the three dimensional annulus $V = \{v \in \mathbb{R}^3 : v_{\min} \leq |v| \leq v_{\max}\}$, where $0 < v_{\min} < v_{\max} < \infty$.

As a *backward equation*, the NTE is written in the form

$$\begin{aligned} \frac{\partial}{\partial t} \psi_t(r, v) &= v \cdot \nabla \psi_t(r, v) - \sigma(r, v) \psi_t(r, v) \\ &+ \sigma_{\mathbf{s}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{s}}(r, v, v') dv' + \sigma_{\mathbf{f}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{f}}(r, v, v') dv', \end{aligned} \quad (3.1)$$

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

- $\sigma_{\mathbf{s}}(r, v)$: the rate at which scattering occurs from incoming velocity v ,
- $\sigma_{\mathbf{f}}(r, v)$: the rate at which fission occurs from incoming velocity v ,
- $\sigma(r, v)$: the sum of the rates $\sigma_{\mathbf{f}} + \sigma_{\mathbf{s}}$ and is known as the total cross section,
- $\pi_{\mathbf{s}}(r, v, v') dv'$: the scattering yield at velocity v' from incoming velocity v ,
satisfying $\int_V \pi_{\mathbf{s}}(r, v, v') dv' = 1$, and
- $\pi_{\mathbf{f}}(r, v, v') dv'$: the neutron yield at velocity v' from fission with incoming velocity v ,
satisfying $\int_V \pi_{\mathbf{f}}(r, v, v') dv' < \infty$.

Some or all of the three assumptions below will be used from time to time in our main results.

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

(H2): We have $\sigma_{\mathbf{s}} \pi_{\mathbf{s}} + \sigma_{\mathbf{f}} \pi_{\mathbf{f}} > 0$ on $D \times V \times V$.

(H3): There is an open ball B compactly embedded in D such that $\sigma_{\mathbf{f}} \pi_{\mathbf{f}} > 0$ on $B \times V \times V$.

It is also usual to insist on the physical boundary conditions

$$\begin{cases} \psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \psi_t(r, v) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } v \cdot \mathbf{n}_r > 0, \end{cases} \quad (3.2)$$

where \mathbf{n}_r is the outward unit normal at $r \in \partial D$ and $g : D \times V \rightarrow [0, \infty)$ is a bounded, measurable function on which we will later impose further conditions. Physically, these boundary conditions mean that any neutron starting on the boundary of the reactor with velocity pointing outwards will be ‘killed’.

Formally speaking, (3.1) as stated is ill defined (because of regularity issues associated to

the transport operator $v \cdot \nabla$) and has traditionally otherwise appeared in applied mathematics and physics literature in the form of an *abstract Cauchy problem* on $L_2(D \times V)$, the space of square integrable functions on $D \times V$. This has formed the principle historical outlook of the analysis of the NTE, appealing to c_0 -semigroup theory; see for example the classical works of [13, 29, 31, 30, 26, 2, 39, 11, 12, 37, 22, 28, 38].

The connection of the NTE via semigroup theory to an underlying stochastic process has, in contrast, received a very limited amount of attention; cf [11, 28, 33]. Accordingly the stochastic analysis of (3.1) has seen very little development in light of recent innovations in the relevant theory of stochastic processes.

In the current article, we are more interested in exploring how NTE can be interpreted as a *mild equation*, describing the mean semigroup evolution of the stochastic process that models the underlying physical process of neutron fission, transport, scatter and absorption. More precisely, we have two main contributions: (i) to develop a new precise statement of the form $\psi_t \sim e^{\lambda_* t} c_g \varphi + o(e^{\lambda_* t})$, where λ_* and φ are a leading eigenvalue and eigenfunction associated to the NTE and c_g is a constant that depends on the initial data g ; (ii) to make the first step in understanding how the growth of the solution to the NTE relative to its lead eigenfunction plays out in terms of the aforementioned physical stochastic process and an associated martingale.

This paper follows the review article [10] which consolidates the existing c_0 -semigroup approach and how it relates to the stochastic representation. A deeper subsequent analysis in the direction of our second objective is continued in the accompanying paper [20]. Further numerical and Monte-Carlo considerations based on our stochastic approach will also appear in forthcoming work [9].

In order to consider the probabilistic perspective, we start by defining the underlying stochastic processes which mimics the physics of neutron fission, transport, scattering and absorption.

3.2 The physical process and the mild NTE

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 finite atomic measures

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

where δ is the Dirac measure, 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 $\mathcal{M}(D \times V) := \{\sum_{i=1}^n \delta_{(r_i, v_i)} : n \in \mathbb{N}, (r_i, v_i) \in D \times V, i = 1, \dots, n\}$ which evolves randomly as follows.

A particle positioned at r with velocity v will continue to move along the trajectory $r + vt$, 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 $T_{\mathfrak{s}}$ for the random time that scattering may occur, then independently of any other physical event that may affect the neutron, $\Pr(T_{\mathfrak{s}} > t) = \exp\{-\int_0^t \sigma_{\mathfrak{s}}(r + vs, v) ds\}$, for $t \geq 0$.

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

- (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 $T_{\mathfrak{f}}$ for the random time that fission may occur, then independently of any other physical event that may affect the neutron, $\Pr(T_{\mathfrak{f}} > t) = \exp\{-\int_0^t \sigma_{\mathfrak{f}}(r + vs, v) ds\}$, for $t \geq 0$.

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 $(v_i : i = 1, \dots, N)$. The outgoing velocities are described by the atomic random measure

$$\mathcal{Z}(A) := \sum_{i=1}^N \delta_{v_i}(A), \quad A \in \mathcal{B}(V). \quad (3.3)$$

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 probabilities $\mathcal{P}_{(r,v)}$ are such that, for $v' \in V$, for bounded and measurable $g : V \rightarrow [0, \infty)$,

$$\int_V g(v') \pi_{\mathfrak{f}}(r, v, v') dv' = \mathcal{E}_{(r,v)} \left[\int_V g(v') \mathcal{Z}(dv') \right] =: \mathcal{E}_{(r,v)}[\langle g, \mathcal{Z} \rangle]. \quad (3.4)$$

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 Markov motion is nothing more than movement in a straight line at velocity v .

Remark 5. It is worth noting how the assumptions (H1)-(H3) play out in the construction of the NBP. Whilst they serve as a sufficient conditions, they are not necessary. For example, one could equally assume that e.g. there are two open domains $B_{\mathfrak{s}}$ and $B_{\mathfrak{f}}$ (which may or may not intersect) contained in D on which $\sigma_{\mathfrak{s}}(r, v) \pi_{\mathfrak{s}}(r, v, v') > 0$, for $B_{\mathfrak{s}} \times V \times V$ and $\sigma_{\mathfrak{f}}(r, v) \pi_{\mathfrak{f}}(r, v, v') > 0$, for $B_{\mathfrak{f}} \times V \times V$, respectively. This would ensure that, at least starting from some configurations

$(r, v) \in D \times V$, the NBP could access regions where scatter or fission occurs with positive probability. From there, the particle system will thus propagate by allowing further opportunities for scatter or fission. That said, there will also be some initial configurations $(r, v) \in D \times V$ for which the particles will neither scatter nor undergo fission and head straight for the boundary ∂D , whereupon they are killed.

This example informally alerts us to the notion of ‘*irreducibility*’ of the state space. For contrast, and to highlight the issue further, it is worth comparing the situation to e.g. a branching Brownian motion on a smooth, convex, bounded domain of $D \subseteq \mathbb{R}^d$ for which the branching rate is supported only on a subdomain B strictly contained in D . In that setting, the Brownian motion of a given particle would always be able to ‘find’ the region B with positive probability, where branching can occur (thus propagating the stochastic process in a non-trivial way). Through this comparison, we see that the piecewise linear spatial paths of neutrons in the NBP, although simpler to depict than the path of a Brownian motion, are significantly more irregular. The assumption (H2) may be thus be thought of as a sufficient condition to ensure irreducibility of the state space by enforcing the possibility of either scatter or fission (but not necessarily the possibility of both), whereas assumption (H3) ensures that there is at least one area of the domain where fission occurs. The condition (H1) simply ensures that activity (scatter and fission) cannot happen too fast, and hence the eventuality of explosion in finite time does not appear in our forthcoming calculations.

Remark 6. The NBP is parameterised by the quantities $\sigma_s, \pi_s, \sigma_f$ 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 $(\sigma_s, \pi_s, \sigma_f, \mathcal{P})$ -NBP. It is associated to the NTE via the relation (3.4), and, although a $(\sigma_s, \pi_s, \sigma_f, \mathcal{P})$ -NBP is uniquely defined, a NBP specified by $(\sigma_s, \pi_s, \sigma_f, \pi_f)$ alone is not.

What is of importance for the purpose of our analysis, however, is that for the given quadruple $(\sigma_s, \pi_s, \sigma_f, \pi_f)$, at least one $(\sigma_s, \pi_s, \sigma_f, \mathcal{P})$ -NBP exists such that (3.4) holds. It is relatively easy to construct an example of \mathcal{P} as such.

Indeed, let us suppose (H1) and (H3) hold. Then, for a given π_f , define

$$n_{\max} = \min\{k \geq 1 : \sup_{(r,v) \in D \times V} \int_V \pi_f(r, v, v') dv' \leq k\}.$$

Define the ensemble $(v_i, i = 1, \dots, N)$ such that: (i) $N \in \{0, n_{\max}\}$; (ii) the probability of the event $\{N = n_{\max}\}$ under $\mathcal{P}_{(r,v)}$ is given by $\int_V \pi_f(r, v, v'') dv'' / n_{\max}$; (iii) on the event $\{N = n_{\max}\}$, each of the n_{\max} neutrons are released with the same velocities $v_1 = \dots = v_{n_{\max}}$; (iv) the distribution of this common velocity is given by

$$\mathcal{P}_{(r,v)}(v_i \in dv' | N = n_{\max}) = \frac{\pi_f(r, v, v')}{\int_V \pi_f(r, v, v'') dv''} dv',$$

for $i = 1, \dots, n_{\max}$.

With the construction (i)-(iv) for $\mathcal{P}_{(r,v)}$, we have for bounded and measurable $g : V \rightarrow$

$[0, \infty)$,

$$\begin{aligned}
& \int_V g(v') \pi_{\mathbf{f}}(r, v, v') dv' \\
&= 0 \times (1 - \mathcal{P}_{(r,v)}(N = n_{\max})) + \mathcal{P}_{(r,v)}(N = n_{\max}) n_{\max} \int_V g(v') \mathcal{P}_{(r,v)}(v_i \in dv' | N = n_{\max}) \\
&= \frac{\int_V \pi_{\mathbf{f}}(r, v, v'') dv''}{n_{\max}} n_{\max} \int_V g(v') \frac{\pi_{\mathbf{f}}(r, v, v')}{\int_V \pi_{\mathbf{f}}(r, v, v'') dv''} dv' \\
&= \int_V g(v') \pi_{\mathbf{f}}(r, v, v') dv',
\end{aligned}$$

thus matching (3.4), as required.

It is interesting to note that the construction above is precisely what happens in industrial models of nuclear reactor cores (for which only the cross-sections $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \pi_{\mathbf{f}})$ are known) when it comes to Monte-Carlo simulation; see further discussion below as well as [9].

The maximum number of neutrons that can be emitted during a fission event with positive probability is finite. 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, albeit, in reality it is more likely that 2 or 3 are released. We will thus occasionally work with:

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

In particular this means that $\sup_{r \in D, v \in V} \int_V \pi_{\mathbf{f}}(r, v, v') dv' \leq n_{\max}$.

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 , each with configuration (r_i, v_i) , $i = 1, \dots, n$.

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 $\mathcal{F}_t := \sigma((r_i(s), v_i(s)) : i = 1, \dots, N_s, s \leq t)$, $t \geq 0$. That is to say, for all bounded and measurable $g : D \times V \rightarrow [0, \infty)$ and $\mu \in \mathcal{M}(D \times V)$ written $\mu = \sum_{i=1}^n \delta_{(r_i, v_i)}$, we have $\mathbb{E}_{\mu}[\prod_{i=1}^{N_t} g(r_i(t), v_i(t))] = \prod_{i=1}^n u_t[g](r_i, v_i)$, for $t \geq 0, r \in D, v \in V$, where $u_t[g](r, v) := \mathbb{E}_{\delta_{(r,v)}}[\prod_{i=1}^{N_t} g(r_i(t), v_i(t))]$. *In this setting it is also customary to work with the notion that the empty product is valued as unity; see [23, 24, 25].*

What is of particular interest to us in the context of the NTE is the *expectation semigroup* of the neutron branching process. More precisely, and with pre-emptive notation, we are interested in

$$\psi_t[g](r, v) := \mathbb{E}_{\delta_{(r,v)}}[\langle g, X_t \rangle], \quad t \geq 0, r \in \bar{D}, v \in V, \quad (3.5)$$

for $g \in L_{\infty}^+(D \times V)$, the space of non-negative 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 (3.4) and written $\langle g, X_t \rangle$ to mean $\int_{D \times V} g(r, v) X_t(dr, dv)$.

To see why $(\psi_t, t \geq 0)$ deserves the name of expectation semigroup, it is a straightforward

exercise with the help of the Markov branching property to show that

$$\psi_{t+s}[g](r, v) = \psi_t[\psi_s[g]](r, v) \quad s, t \geq 0. \quad (3.6)$$

The connection of the expectation semigroup (3.5) with the NTE (3.1) was explored in the recent article [10] (see also older work in [11, 28]). In order to present the relevant findings, let us momentarily introduce some notation. The deterministic evolution $\mathcal{U}_t[g](r, v) = g(r + vt, v)\mathbf{1}_{\{t < \kappa_{r,v}^D\}}$, $t \geq 0$, and $\kappa_{r,v}^D := \inf\{t > 0 : r + vt \notin D\}$, represents the advection semigroup associated with a single neutron travelling at velocity v from r . The backwards scatter operator is denoted by

$$\overleftarrow{\mathbb{S}}f(r, v) = \sigma_{\mathfrak{s}}(r, v) \int_V f(r, v')\pi_{\mathfrak{s}}(r, v, v')dv' - \sigma_{\mathfrak{s}}(r, v)f(r, v) \quad (3.7)$$

and the backwards fission operator is given by

$$\overleftarrow{\mathbb{F}}f(r, v) = \sigma_{\mathfrak{f}}(r, v) \int_V f(r, v')\pi_{\mathfrak{f}}(r, v, v')dv' - \sigma_{\mathfrak{f}}(r, v)f(r, v), \quad (3.8)$$

for $f \in L_{\infty}^+(D \times V)$, such that both $\overleftarrow{\mathbb{S}}$ and $\overleftarrow{\mathbb{F}}$ are defined on $D \times V$ and zero otherwise.

Lemma 3 ([10]). *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 (3.5), 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 to the so-called mild equation (also called a Duhamel solution in the PDE literature):*

$$\psi_t[g] = \mathcal{U}_t[g] + \int_0^t \mathcal{U}_s[(\overleftarrow{\mathbb{S}} + \overleftarrow{\mathbb{F}})\psi_{t-s}[g]]ds, \quad t \geq 0, \quad (3.9)$$

for which (3.2) holds.

The fact that (3.5) solves (3.9) is a simple matter of conditioning the expression in (3.5) on the first fission or scatter event (whichever occurs first) and rearranging the resulting equation. Uniqueness is a matter of working in the right way with Grönwall's Lemma. The association of (3.9) with (3.1) in this way was also explored in Theorem 7.1 [10], where it was shown that the unique solution to (3.1) when seen as an abstract Cauchy problem on $L_2(D \times V)$ agrees with the unique solution to (3.9) in the $L_2(D \times V)$ norm.

The reader should note that we do not need (H3) or (H4) as the result does not require information about the pathwise behaviour of any associated underlying stochastic processes. Nor does it distinguish between the settings that $\overleftarrow{\mathbb{F}}$ is present or not.

3.3 Perron-Frobenius asymptotics

As alluded to above, one of the classical ways in which neutron flux is understood is to look for the leading eigenvalue and associated ground state eigenfunction. Roughly speaking, this means

looking for an associated triple of eigenvalue $\lambda_* \in \mathbb{R}$, positive right eigenfunction $\varphi : D \times V \rightarrow [0, \infty)$, a left eigenmeasure $\tilde{\varphi}(r, v)drdv$ on $D \times V$ in $L_2^+(D \times V)$ (the cone of non-negative square integrable functions on $D \times V$) such that $\langle f, \psi_t[\varphi] \rangle = e^{\lambda_* t} \langle f, \varphi \rangle$ and $\langle \tilde{\varphi}, \psi_t[f] \rangle = e^{\lambda_* t} \langle \tilde{\varphi}, f \rangle$, for $t \geq 0$. Here, we again abuse our notation (see the use of $\langle \cdot, \cdot \rangle$ in (3.4) and (3.5)) and write, for $f, g \in L_2^+(D \times V)$, $\langle f, g \rangle = \int_{D \times V} f(r, v)g(r, v)drdv$. With the eigentriple in hand, it is a common point of analysis that, to leading order, the NTE (3.1) is solved through the approximation

$$\psi_t(r, v) = e^{\lambda_* t} \langle g, \tilde{\varphi} \rangle \varphi(r, v) + o(e^{\lambda_* t}), \quad t \geq 0, r \in D, v \in V, \quad (3.10)$$

where the sense of the equality depends on how one interprets the NTE (i.e. as an abstract Cauchy problem or in its mild form).

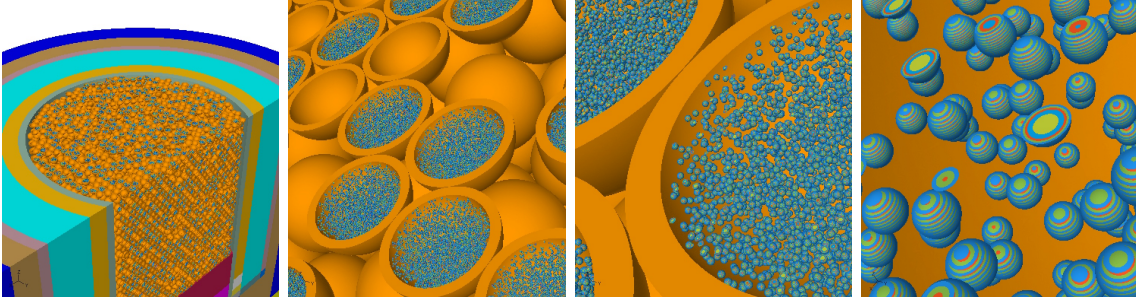


Figure 3-1: Zooming into a virtual model of a pebble bed nuclear reactor core made from tennis ball sized orbs which encapsulate uranium pellets. From left to right, the diagrams illustrate detail from metres down to millimetres. Colour indicates the different regions in which the cross sections $\sigma_s, \sigma_f, \pi_s, \pi_f$ are constant. The structural design of such a reactor can easily be stored as a virtual environment (i.e. storing the coordinates of the different geometrical domains and the material properties in each domain) with just circa 150MB of data, on to which extensive data libraries of industrial numerical values for the respective quantities $\sigma_s, \sigma_f, \pi_s, \pi_f$ can be mapped.

The eigenfunction $\tilde{\varphi}$ is called the *importance map* and offers a quasi-stationary profile of radioactive activity (unless $\lambda_* = 0$, in which case it is a stationary profile). Indeed, in modern nuclear reactor design and safety regulation, it is usually the case that virtual reactor models such as the one seen in Figure 3-1 (an example of a uranium pebble bed reactor) are designed such that $\lambda_* = 0$ and the behaviour of $\tilde{\varphi}$ within spatial domains on the human scale remains 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 $\tilde{\varphi}(r, v)drdv$. Giving a sensible meaning to (3.10) will play an important part in unraveling the analysis of stochastic representations of solutions to the NTE as well. Moreover, in additional forthcoming work [9], we will also see that our asymptotic (3.10), together with the accompanying stochastic analysis developed here, has influence on a number of completely new Monte Carlo methods associated with the NTE that, in turn, bears relevance to the applied NTE literature.

The approximation (3.10) can be seen as a functional version of the Perron-Frobenius

Theorem, in particular when noting via (3.9) that we can understand $\psi_t[g]$ as a semigroup. Many attempts have been made to generalise the notion of the Perron-Frobenius decomposition to semigroups of Markov processes with countable and uncountable state spaces, as well as with killing and mass creation (see for example [14, 34, 35, 36]), using what has come to be known as R -theory. The conditions there seem difficult to verify in the current setting.

More recently, [7, 8] have provided an alternative approach to the R -theory presented in aforementioned works. In the current context, Theorem 2.1 and Proposition 2.3 of [7] will help us to achieve the global result, given below. To state it we need to introduce the quantity

$$\alpha(r, v)\pi(r, v, v') = \sigma_{\mathfrak{s}}(r, v)\pi_{\mathfrak{s}}(r, v, v') + \sigma_{\mathfrak{f}}(r, v)\pi_{\mathfrak{f}}(r, v, v') \quad r \in D, v, v' \in V, \quad (3.11)$$

where π is taken to be a probability density. As such it necessarily follows that

$$\alpha(r, v) = \sigma_{\mathfrak{s}}(r, v) + \sigma_{\mathfrak{f}}(r, v) \int_V \pi_{\mathfrak{f}}(r, v, v') dv'. \quad (3.12)$$

Theorem 7. *Suppose that (H1) holds as well as **(H2)***: $\inf_{r \in D, v, v' \in V} \alpha(r, v)\pi(r, v, v') > 0$.*

Then, for the semigroup $(\psi_t, t \geq 0)$ identified by (3.9), there exists a $\lambda_ \in \mathbb{R}$, a positive⁴ right eigenfunction $\varphi \in L_{\infty}^+(D \times V)$ and a left eigenmeasure which is absolutely continuous with respect to Lebesgue measure on $D \times V$ with density $\tilde{\varphi} \in L_{\infty}^+(D \times V)$, both having associated eigenvalue $e^{\lambda_* t}$, and such that φ (resp. $\tilde{\varphi}$) is uniformly (resp. a.e. uniformly) bounded away from zero on each compactly embedded subset of $D \times V$. In particular, for all $g \in L_{\infty}^+(D \times V)$,*

$$\langle \tilde{\varphi}, \psi_t[g] \rangle = e^{\lambda_* t} \langle \tilde{\varphi}, g \rangle \quad (\text{resp. } \psi_t[\varphi] = e^{\lambda_* t} \varphi) \quad t \geq 0. \quad (3.13)$$

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

$$\sup_{g \in L_{\infty}^+(D \times V): \|g\|_{\infty} \leq 1} \left\| e^{-\lambda_* t} \varphi^{-1} \psi_t[g] - \langle \tilde{\varphi}, g \rangle \right\|_{\infty} = O(e^{-\varepsilon t}) \text{ as } t \rightarrow +\infty. \quad (3.14)$$

This result differs significantly from what is already in the literature principally through the assumptions on the cross-sections, the strict positivity properties and the uniform boundedness of $\varphi, \tilde{\varphi}$ and uniformity in the mode of convergence. In existing literature (3.14) is usually given in the L_p setting, where $1 < p < \infty$ is strictly enforced due to the nature of the c_0 -semigroup perturbation analysis; cf. [12, 37] and the discussion in [10].

The proof of Theorem 7 is a non-trivial application of the recent theory of [7, 8] in that verifying their assumptions (which essentially leads to the full statement of Theorem 7) is highly technical, taking account of the dimension of the system and the piecewise linear (and hence irregular) nature of the neutron paths in the underlying NBP.

Once again the assumptions (H3) and (H4) are unnecessary. As we shall shortly see, the

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

result relies on the treatment of the sum of the operators $\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}}$ as a single object, re-written as a scattering generator with action

$$\overleftarrow{\mathbf{S}}' f(r, v) = \int_V (f(r, v) - f(r, v')) \alpha(r, v) \pi(r, v, v') dv', \quad r \in D, v \in V.$$

The assumption (H2)* is a condition on the intensity of this new generator. In this sense, the need for fission or for control of the pathwise behaviour of number of offspring (other than through their mean) is not needed.

3.4 Neutron random walk and many-to-one methodology

There is a second stochastic representation of the unique solution to (3.9) which will form the basis of our proof of Theorem 7. 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, $\zeta(r, v)$, $r \in D, v \in V$, and scatter probability densities $\varpi(r, v, v')$, $r \in D, v, v' \in V$ such that $\int_V \varpi(r, v, v') dv' = 1$ for all $r \in D, v \in V$. Simply, when issued from r 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 $T_{\mathbf{s}}$ a scattering occurs, where $\Pr(T_{\mathbf{s}} > t) = \exp\{-\int_0^t \zeta(r + vs, v) ds\}$, for $t \geq 0$. When the scattering event occurs in position-velocity configuration (r, v) , a new velocity v' is selected with probability $\varpi(r, v, v') dv'$. 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, neither R nor Υ alone is Markovian. We call the process (R, Υ) an $\zeta\varpi$ -NRW. It is worth remarking that when $\zeta\varpi$ is given as a single rate function, the density ϖ , and hence the rate ζ , is uniquely identified by normalising of the given product form $\zeta\varpi$ to make it a probability distribution.

To describe the second stochastic representation of (3.9), we define

$$\beta(r, v) = \sigma_{\mathbf{f}}(r, v) \left(\int_V \pi_{\mathbf{f}}(r, v, v') dv' - 1 \right) \geq - \sup_{r \in D, v \in V} \sigma_{\mathbf{f}}(r, v) > -\infty, \quad (3.15)$$

where the lower bound is due to assumption (H1). The following result was established in Lemma 7.1 of [10].

Lemma 4 (Many-to-one formula, [10]). *Under the assumptions of Lemma 3, we have the second representation*

$$\psi_t[g](r, v) = \mathbf{E}_{(r, v)} \left[e^{\int_0^t \beta(R_s, \Upsilon_s) ds} g(R_t, \Upsilon_t) \mathbf{1}_{\{t < \tau^D\}} \right], \quad t \geq 0, r \in D, v \in V, \quad (3.16)$$

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

Noting that $\bar{\beta} := \sup_{r \in D, v \in V} \beta(r, v) < \infty$ thanks to (H1), let us introduce $\mathbf{P}^\dagger := (\mathbf{P}_t^\dagger, t \geq 0)$ for the expectation semigroup of the $\alpha\pi$ -neutron random walk with potential β , such as is represented by the semigroup (3.16), but now killed at rate $(\bar{\beta} - \beta)$. More precisely, for $g \in L_\infty^+(D \times V)$,

$$\begin{aligned} \mathbf{P}_t^\dagger[g](r, v) &= \psi_t[g](r, v)e^{-\bar{\beta}t} \\ &= \mathbf{E}_{(r, v)} \left[e^{\int_0^t (\beta(R_s, \Upsilon_s) - \bar{\beta}) ds} g(R_t, \Upsilon_t) \mathbf{1}_{\{t < \tau^D\}} \right] \\ &= \mathbf{E}_{(r, v)} \left[g(R_t, \Upsilon_t) \mathbf{1}_{\{t < \mathbf{k}\}} \right] \\ &=: \mathbf{E}_{(r, v)}^\dagger [g(R_t, \Upsilon_t)], \quad t \geq 0, r \in D, v \in V, \end{aligned} \tag{3.17}$$

where

$$\mathbf{k} = \inf\{t > 0 : \int_0^t (\bar{\beta} - \beta(R_s, \Upsilon_s)) ds > \mathbf{e}\} \wedge \tau^D, \tag{3.18}$$

and \mathbf{e} is an independent exponentially distributed random variable with mean 1.

We will naturally write $\mathbf{P}_{(r, v)}^\dagger$ for the (sub)probability measure associated to $\mathbf{E}_{(r, v)}^\dagger$, $r \in \bar{D}, v \in V$. The family $\mathbf{P}^\dagger := (\mathbf{P}_{(r, v)}^\dagger, r \in \bar{D}, v \in V)$ now defines a Markov family of probability measures on the path space of the neutron random walk with cemetery state $\{\dagger\}$, which is where the path is sent when hitting the boundary ∂D or the clock associated to the killing rate $\bar{\beta} - \beta$ rings. *We note for future calculations that we can extend the domain of functions on $D \times V$ to accommodate taking a value on $\{\dagger\}$ by insisting that this value is always 0.*

Our strategy for proving Theorem 7 thus boils down to understanding the evolution of the semigroup of the NRW $((R, \Upsilon), \mathbf{P}^\dagger)$. In this sense, we see that the essence of Theorem 7 is, roughly speaking, a classical Perron-Frobenius-type problem for the semigroup of a Markov process; namely $((R, \Upsilon), \mathbf{P}^\dagger)$.

It is also worthy of note that, given the role the $\alpha\pi$ -NRW in the proof of Theorem 7, we can also interpret the role of the assumptions (H1) and (H2)* in terms of this process. The condition (H1) ensures that scattering cannot occur too fast. We can describe the condition (H2)* by saying that it offers ‘strong irreducibility’ of the $\alpha\pi$ -NRW (where e.g. we could say that (H2) only offers ‘weak irreducibility’).

As alluded to previously, we can also see why the absence of the assumption (H3) is not a problem. In the event that e.g. $\sigma_{\mathbf{f}}\pi_{\mathbf{f}}$ is identically zero, the original NBP is nothing more than a $\sigma_{\mathbf{s}}\pi_{\mathbf{s}}$ -NRW, i.e. $\alpha\pi = \sigma_{\mathbf{s}}\pi_{\mathbf{s}}$ and $\beta \leq 0$. As such the analysis in the proof of Theorem 7, which fundamentally concerns a NRW with a ‘strictly irreducible’ state space and killing is still valid. Similarly, the inclusion of (H4) is not necessary as we only need control over the kernel $\alpha\pi$ for the purpose of analysing the associated NRW and not the pathwise behaviour of the otherwise associated NBP.

3.5 The ground state martingale

As an application of the Perron-Frobenius behaviour of the linear semigroups discussed in Theorem 7, we complete the summary of the main results of this paper by discussing how the existence of the right eigenfunction φ plays directly into the existence of a classical (ground state) martingale for the underlying physical process. Analogues of this martingale appear in the setting of all spatial branching processes and are sometimes referred to as ‘the additive martingale’ (see for example the recent monograph [40] which discusses the analogous setting for branching random walks, or [3] for fragmentation processes).

Under the assumptions of Theorem 7, thanks to the semigroup property of (3.5) and the invariance of φ in Theorem 7, it is now easy to see that

$$W_t := e^{-\lambda_* t} \frac{\langle \varphi, X_t \rangle}{\langle \varphi, \mu \rangle}, \quad t \geq 0, \quad (3.19)$$

is a unit mean martingale under \mathbb{P}_μ where $\mu \in \mathcal{M}(D \times V)$. It is worth noting that this claim is not so easy to make under analogues of Theorem 7 found in previous literature (cf. [11, 12, 37]) as the setting of the eigenfunction φ in an L_p space would make it difficult to make sense of expectations of the inner product $\langle \varphi, X_t \rangle$ without saying more about the mean semigroup of $(X_t, t \geq 0)$.

As a non-negative martingale, the almost sure limit of (3.19) is assured. Our second main result tells us precisely when this martingale limit is non-zero. Before stating the theorem, we require one more assumption on the fission rate and kernel, which is a stricter version of (H3).

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

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

Theorem 8. *For the $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \mathcal{P})$ -NBP satisfying (H1), (H2)* and (H4), we have the following cases for the martingale $W = (W_t, t \geq 0)$:*

- (i) *If $\lambda_* > 0$ and (H3) holds, then W is $L_1(\mathbb{P})$ convergent;*
- (ii) *If $\lambda_* < 0$ and (H3) holds, then $W_\infty = 0$ almost surely;*
- (iii) *If $\lambda_* = 0$ and (H3)* holds, then $W_\infty = 0$ almost surely.*

As we can see from the above theorem, the critical case requires slightly more stringent conditions than the super- or sub-critical cases. However, if we assume the conditions of the critical case across the board, we get the aesthetically more pleasing corollary below.

Corollary 1. *For the $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \mathcal{P})$ -NBP satisfying (H1), (H2)*, (H3)* and (H4), the martingale W is $L_1(\mathbb{P})$ convergent if and only if $\lambda_* > 0$ and otherwise $W_\infty = 0$, almost surely.*

Note that, unlike many spatial branching process (e.g. the classical result of [6]), there is no ‘*xlogx*’ condition thanks to the assumption (H4) and a precise dichotomy on λ_* emerges. The result mimics a behavioural trait that has been observed for branching diffusions in compact domains in e.g. [18]. In essence it states that in the competing physical processes of fission, transport, scattering and absorption, it is the lead eigenvalue which dictates growth or decay of mass. In this respect we can also mimic other similar results in the spatial branching process literature (cf. [1, 19, 27]), the proof of which falls out of the proof of Theorem 8.

Corollary 2. *For the $(\sigma_{\mathfrak{s}}, \pi_{\mathfrak{s}}, \sigma_{\mathfrak{f}}, \mathcal{P})$ -NBP satisfying the assumptions (H1), (H2)*, (H3) and (H4), when $\lambda_* > 0$, the martingale $(W_t, t \geq 0)$ is $L_2(\mathbb{P})$ convergent.*

It is particularly interesting to note that in the setting of a critical system, $\lambda_* = 0$, which is typically what is envisaged for a nuclear reactor, the above results evidence the hypothesis that the fission process eventually dies out (similarly to other examples of critical branching processes). To verify this rigorously, one needs an almost sure growth result for the particle system which would take the format

$$\lim_{t \rightarrow \infty} e^{-\lambda_* t} \frac{\langle g, X_t \rangle}{\langle \varphi, \mu \rangle} = \langle g, \tilde{\varphi} \rangle W_\infty, \quad (3.20)$$

\mathbb{P}_μ -almost surely, for all $g \in L_\infty^+(D \times V)$. This is a much more difficult result than the one stated in Theorem 8 and is addressed in a second instalment to this paper; see [20]. The reader should note that (3.20) verifies what has been known in the nuclear industry for a long time. Namely that critical nuclear reactors will not persist in energy generation, but will eventually cease working, corresponding to the case that $W_\infty = 0$.

3.6 Neutron random walk and spine decomposition

As with many spatial branching processes, the most efficient way to analyse martingale convergence is through the pathwise behaviour of the particle system (known as a *spine decomposition*) when considered under a change of measure induced by the martingale itself. Whilst classical in the branching process literature, it is unknown in the setting of neutron transport. We will devote the remainder of this section to describing the pathwise spine decomposition of the physical process, our final main contribution.

We are interested in the change of measure

$$\left. \frac{d\mathbb{P}_\mu^\varphi}{d\mathbb{P}_\mu} \right|_{\mathcal{F}_t} = W_t, \quad t \geq 0, \quad (3.21)$$

for the NBP with characteristics $\sigma_{\mathfrak{s}}, \pi_{\mathfrak{s}}, \sigma_{\mathfrak{f}}, \mathcal{P}$ (cf. Remark 6), where μ belongs to the space of finite atomic measures $\mathcal{M}(D \times V)$.

In the next theorem we will formalise an understanding of this change of measure in terms

of another $\mathcal{M}(D \times V)$ -valued stochastic process

$$X^\varphi := (X_t^\varphi, t \geq 0) \text{ with probabilities } \tilde{\mathbb{P}}^\varphi := (\tilde{\mathbb{P}}_\mu^\varphi, \mu \in \mathcal{M}(D \times V)), \quad (3.22)$$

which we will now describe through an algorithmic construction.

1. From the initial configuration $\mu \in \mathcal{M}(D \times V)$ with an arbitrary enumeration of particles, the i -th neutron is selected and marked ‘*spine*’ with empirical probability

$$\frac{\varphi(r_i, v_i)}{\langle \varphi, \mu \rangle}.$$

2. The neutrons $j \neq i$ in the initial configuration that are not marked ‘*spine*’, each issue independent copies of $(X, \mathbb{P}_{\delta_{(r_j, v_j)}})$ respectively.

3. For the marked neutron, issue a NRW characterised by the rate function

$$\sigma_s(r, v) \frac{\varphi(r, v')}{\varphi(r, v)} \pi_s(r, v, v'), \quad r \in D, v, v' \in V.$$

4. The marked neutron undergoes fission at the accelerated rate $\varphi(r, v)^{-1}(\bar{\mathbb{F}} + \sigma_{\mathbf{f}}\mathbf{I})\varphi(r, v)$, when in physical configuration $(r, v) \in D \times V$, at which point, it scatters a random number of particles according to the random measure on V given by $(\mathcal{Z}, \mathcal{P}_{(r, v)}^\varphi)$ where

$$\frac{d\mathcal{P}_{(r, v)}^\varphi}{d\mathcal{P}_{(r, v)}} = \frac{\langle \varphi, \mathcal{Z} \rangle}{\mathcal{E}_{(r, v)}[\langle \varphi, \mathcal{Z} \rangle]}. \quad (3.23)$$

5. When fission of the marked neutron occurs in physical configuration $(r, v) \in D \times V$, set

$$\mu = \sum_{i=1}^n \delta_{(r, v_i)}, \text{ where, in the previous step, } \mathcal{Z} = \sum_{i=1}^n \delta_{v_i},$$

and repeat step 1.

The process X_t^φ describes the physical configuration (position and velocity) of all the particles in the system at time t , for $t \geq 0$ (i.e. ignoring the marked genealogy). We will also be interested in the configuration of the single genealogical line of descent which has been marked ‘*spine*’. This process, referred to simply as *the spine*, will be denoted by $(R^\varphi, \Upsilon^\varphi) := ((R_t^\varphi, \Upsilon_t^\varphi), t \geq 0)$. Together, the processes $(X^\varphi, (R^\varphi, \Upsilon^\varphi))$ make a Markov pair, whose probabilities we will denote by $(\tilde{\mathbb{P}}_{\mu, (r, v)}^\varphi, \mu \in \mathcal{M}(D \times V), (r, v) \in V \times D)$. Note in particular that

$$\tilde{\mathbb{P}}_\mu^\varphi = \sum_{i=1}^n \frac{\varphi(r_i, v_i)}{\langle \varphi, \mu \rangle} \tilde{\mathbb{P}}_{\mu, (r_i, v_i)}^\varphi$$

when $\mu = \sum_{i=1}^n \delta_{(r_i, v_i)}$.

Theorem 9. *Under assumptions (H1), (H2) and (H4), the process $(X^\varphi, \tilde{\mathbb{P}}^\varphi)$ is Markovian and equal in law to (X, \mathbb{P}^φ) , where $\mathbb{P}^\varphi = (\mathbb{P}_\mu^\varphi, \mu \in \mathcal{M}(D \times V))$.*

It is also worth understanding the dynamics of the spine $(R^\varphi, \Upsilon^\varphi)$. For convenience, let us denote the family of probabilities of the latter by $\tilde{\mathbf{P}}^\varphi = (\tilde{\mathbf{P}}_{(r,v)}^\varphi, (r, v) \in D \times V)$, in other words, the marginals of $(\tilde{\mathbb{P}}_{\mu, (r,v)}^\varphi, \mu \in \mathcal{M}(D \times V), (r, v) \in V \times D)$.

Next we define the probabilities $\mathbf{P}^\varphi := (\mathbf{P}_{(r,v)}^\varphi, (r, v) \in D \times V)$ to describe the law of an $\alpha^\varphi \pi^\varphi$ -NRW, where

$$\alpha^\varphi(r, v) \pi^\varphi(r, v, v') = \frac{\varphi(r, v')}{\varphi(r, v)} (\sigma_{\mathbf{s}}(r, v) \pi_{\mathbf{s}}(r, v, v') + \sigma_{\mathbf{f}}(r, v) \pi_{\mathbf{f}}(r, v, v')), \quad (3.24)$$

for $r \in D, v, v' \in V$. We are now ready to identify the spine.

Lemma 5. *Under assumptions (H1), (H2) and (H4), the process $((R^\varphi, \Upsilon^\varphi), \tilde{\mathbf{P}}^\varphi)$ is a NRW equal in law to $((R, \Upsilon), \mathbf{P}^\varphi)$ and, moreover,*

$$\left. \frac{d\mathbf{P}_{(r,v)}^\varphi}{d\mathbf{P}_{(r,v)}} \right|_{\mathcal{F}_t} = e^{-\lambda_* t + \int_0^t \beta(R_s, \Upsilon_s) ds} \frac{\varphi(R_t, \Upsilon_t)}{\varphi(r, v)} \mathbf{1}_{(t < \tau^D)}, \quad t \geq 0, r \in D, v \in V, \quad (3.25)$$

from which we deduce that $((R, \Upsilon), \mathbf{P}^\varphi)$ is conservative with a stationary distribution $\varphi \tilde{\varphi}(r, v) dr dv$ on $D \times V$. (Recall that (R, Υ) under \mathbf{P} is the $\alpha \pi$ -NRW that appears in the many-to-one Lemma 4.)

Now that we have stated all of our main results, it is worth noting that, in places, the analysis echoes very similar issues that have very recently appeared in the analysis of growth-fragmentation equations, see e.g. [5] and [4], and for good reason. Growth-fragmentation equations, although dealing with a particle system in which particles' mass is positive-valued and for which there is no consideration of classical 'velocity', the dynamics of fragmentation shares the phenomenon of non-local branching. This explains the appearance of integral operators. Moreover, a combination of Lévy-type and piecewise linear movement of particles in the growth-fragmentation setting also mirrors the phenomenon of advection and scattering in the NTE and the associated operators.

In the rest of the paper we prove Theorem 7, Theorem 9, Lemma 5, Theorem 8 and Corollary 2 in that order.

3.7 Proof of Theorem 7

Our approach to proving Theorem 7 will be to extract the existence of the eigentriple λ_* , φ and $\tilde{\varphi}$ for the expectation semigroup $(\psi_t, t \geq 0)$ from the existence of a similar triple of the semigroup $(\mathbf{P}_t^\dagger, t \geq 0)$ defined in (3.17). Indeed, from (3.17), it is clear that when the latter exists, the eigenfunctions of the former are the same and the eigenvalues differ only by the constant $\bar{\beta}$.

Throughout this section, we assume the assumptions of Theorem 7 are in force.

As alluded to earlier, what lies at the core of our proof is the general result of Theorem 2.1 and Proposition 2.3 of [7] and Theorem 2.1 and the discussion around (1.5) of [8], which, combined in the current context, reads as follows.

Theorem 10. *Suppose that there exists a probability measure ν on $D \times V$ such that*

(A1) *there exist $t_0, c_1 > 0$ such that for each $(r, v) \in D \times V$,*

$$\mathbf{P}_{(r,v)}((R_{t_0}, \Upsilon_{t_0}) \in \cdot \mid t_0 < \mathbf{k}) \geq c_1 \nu(\cdot);$$

(A2) *there exists a constant $c_2 > 0$ such that for each $(r, v) \in D \times V$ and for every $t \geq 0$,*

$$\mathbf{P}_\nu(t < \mathbf{k}) \geq c_2 \mathbf{P}_{(r,v)}(t < \mathbf{k}),$$

where \mathbf{k} was defined in (3.18). Then, there exists $\lambda_c < 0$ such that, there exists an eigenmeasure η on $D \times V$ and a positive right eigenfunction φ of \mathbf{P}^\dagger with eigenvalue $e^{\lambda_c t}$, such that η is a probability measure and $\varphi \in L_\infty^+(D \times V)$, i.e. for all $g \in L_\infty(D \times V)$

$$\eta[\mathbf{P}_t^\dagger[g]] = e^{\lambda_c t} \eta[g] \quad \text{and} \quad \mathbf{P}_t^\dagger[\varphi] = e^{\lambda_c t} \varphi \quad t \geq 0. \quad (3.26)$$

Moreover, there exist $C, \varepsilon > 0$ such that, for $g \in L_\infty^+(D \times V)$ and t sufficiently large (which does not depend on g),

$$\left\| e^{-\lambda_c t} \varphi^{-1} \mathbf{P}_t^\dagger[g] - \eta[g] \right\|_\infty \leq C e^{-\varepsilon t} \|g\|_\infty. \quad (3.27)$$

In particular, setting $g \equiv 1$, as $t \rightarrow \infty$,

$$\left\| e^{-\lambda_c t} \varphi^{-1} \mathbf{P}_t(t < \mathbf{k}) - 1 \right\|_\infty \leq C e^{-\varepsilon t}. \quad (3.28)$$

We aim to prove that assumptions (A1) and (A2) are satisfied, so that the conclusions of the above theorem hold. Then we prove that φ is uniformly bounded away from 0 on each compactly embedded subset of $D \times V$ and that η admits a positive bounded density with respect to the Lebesgue measure on $D \times V$ (see Lemma 9), which concludes the proof of Theorem 7. In order to do so, we start by introducing two alternative assumptions to (A1) and (A2):

There exists an $\varepsilon > 0$ such that

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

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

It is easy to verify that (B1) and (B2) are implied when we assume that D is a non-empty convex domain, as we have done in the introduction. They are also satisfied if, for example, the boundary of D is a smooth, connected, compact manifold and ε is sufficiently

small. Geometrically, (B2) means that each of the sets

$$L_r := \left\{ z \in \mathbb{R}^3 : \frac{\|z - r\|}{\|v\|} \in [s_\varepsilon, t_\varepsilon], v \in K_r \right\}, \quad r \in D \setminus D_\varepsilon \quad (3.29)$$

is included in D_ε and has Lebesgue measure at least $\gamma(t_\varepsilon^2 - s_\varepsilon^2)/2$. Roughly speaking, for each $r \in D$ which is within $\varepsilon \mathbf{v}_{\max}$ of the boundary ∂D , L_r is the set of points from which one can issue a neutron with a velocity chosen from $v \in K_r$ such that (ignoring scattering and fission) we can ensure that it passes through $D \setminus D_\varepsilon$ during the time interval $[s_\varepsilon, t_\varepsilon]$.

Our proof of Theorem 7 thus consists of proving that assumptions (B1) and (B2) imply assumptions (A1) and (A2). Our method is motivated by [7, Section 4.2]; however, we note that our approach accommodates for the more general setting we have here (e.g. $V \subset \mathbb{R}^3$ is bounded and $d = 3$) at the cost of greater technicalities.

We begin by considering several technical lemmas. The first is a straightforward consequence of D being a bounded subset of \mathbb{R}^3 .

Lemma 6. *Let $B(r, v)$ be the ball in \mathbb{R}^3 centred at r with radius v .*

(i) *There exists an integer $\mathbf{n} \geq 1$ and $r_1, \dots, r_n \in D_\varepsilon$ such that $D_\varepsilon \subset \bigcup_{i=1}^{\mathbf{n}} B(r_i, \mathbf{v}_{\max}\varepsilon/32)$ and $D_\varepsilon \cap B(r_i, \mathbf{v}_{\max}\varepsilon/32) \neq \emptyset$ for each $i \in \{1, \dots, \mathbf{n}\}$.*

(ii) *For all $r, r' \in D_\varepsilon$, there exists $m \leq \mathbf{n}$ and i_1, \dots, i_m distinct in $\{1, \dots, \mathbf{n}\}$ such that $r \in B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32)$, $r' \in B(r_{i_m}, \mathbf{v}_{\max}\varepsilon/32)$ and for all $1 \leq j \leq m - 1$, $B(r_{i_j}, \mathbf{v}_{\max}\varepsilon/32) \cap B(r_{i_{j+1}}, \mathbf{v}_{\max}\varepsilon/32) \neq \emptyset$.*

Heuristically, the above lemma ensures that there is a universal covering of D_ε by the balls $B(r_i, \mathbf{v}_{\max}\varepsilon/32)$, $1 \leq i \leq \mathbf{n}$ such that between any two points $r, r' \in D_\varepsilon$, there is a sequence of overlapping balls $B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32), \dots, B(r_{i_m}, \mathbf{v}_{\max}\varepsilon/32)$ that one may pass through in order to get from r to r' .

The next lemma provides a minorization of the law of (R_t, Υ_t) under \mathbf{P}^\dagger . The result is similar to [7, Lemma 4.5]; however, we provide a less geometrical proof by considering a change of variables from Cartesian to polar coordinates. In the statement of the lemma, we use $\text{dist}(r, \partial D)$ for the distance of r from the boundary ∂D .

Define $\underline{\alpha} = \inf_{r \in D, v \in V} \alpha(r, v) > 0$ and $\underline{\pi} = \inf_{r \in D, v, v' \in V} \pi(r, v, v')$. We will also similarly write $\bar{\alpha}$ and $\bar{\pi}$ with obvious meanings. We note that due to the assumption (H1) we have $\bar{\alpha} < \infty$ and $\bar{\pi} < \infty$ and hence, combining this with (H2)* it follows that,

$$\underline{\alpha} = \frac{1}{\bar{\pi}} \inf_{r \in D, v \in V} \alpha(r, v) \bar{\pi} \geq \frac{1}{\bar{\pi}} \inf_{r \in D, v, v' \in V} \alpha(r, v) \pi(r, v, v') > 0,$$

and a similar calculation shows that $\underline{\pi} > 0$.

Lemma 7. *For all $r \in D$, $v \in V$ and $t > 0$ such that $\mathbf{v}_{\max} t < \text{dist}(r, \partial D)$, the law of (R_t, Υ_t)*

under $\mathbf{P}_{(r,v)}^\dagger$, defined in (3.17), satisfies

$$\begin{aligned} \mathbf{P}_{(r,v)}^\dagger(R_t \in dz, \Upsilon_t \in dv) &\geq \frac{Ce^{-\bar{\alpha}t}}{t^2} \left[\frac{t}{2} \left(v_{\max}^2 - \left(v_{\min} \vee \frac{|z-r|}{t} \right)^2 \right) \right. \\ &\quad \left. - |z-r| \left(v_{\max} - v_{\min} \vee \frac{|z-r|}{t} \right) \right] \mathbf{1}_{\{z \in B(r, v_{\max}t)\}} dz dv, \end{aligned} \quad (3.30)$$

where $C > 0$ is a positive constant.

Proof. Fix $r_0 \in D$. Let J_k denote the k^{th} jump time of (R_t, Υ_t) under $\mathbf{P}_{(r,v)}^\dagger$ and let Υ_0 be uniformly distributed on V . Assuming that $v_{\max}t < \text{dist}(r_0, \partial D)$, we first give a minorization of the density of (R_t, Υ_t) , with initial configuration (r_0, Υ_0) , on the event $\{J_1 \leq t < J_2\}$. Note that, on this event, we have

$$R_t = r_0 + J_1 \Upsilon_0 + (t - J_1) \Upsilon_{J_1},$$

where Υ_{J_1} is the velocity of the process after the first jump. Then

$$\begin{aligned} \mathbf{E}_{(r_0, \Upsilon_0)}^\dagger[f(R_t, \Upsilon_t) \mathbf{1}_{\{J_1 \leq t < J_2\}}] &= \int_0^t ds \int_V dv_0 \int_V dv_1 \alpha(r_0 + v_0s, v_0) e^{-\int_0^s \alpha(r_0 + v_0u, v_0) du} e^{-\int_0^{t-s} \alpha(r_0 + v_0s + v_1u, v_1) du} \\ &\quad \times \pi(r_0 + v_0s, v_0, v_1) f(r_0 + v_0s + (t-s)v_1, v_1) \\ &\geq \underline{\alpha} e^{-\bar{\alpha}t} \underline{\pi} \int_V dv_1 \int_0^t ds \int_V dv_0 f(r_0 + sv_0 + (t-s)v_1, v_1), \end{aligned} \quad (3.31)$$

where we have used the bounds on α and π . We now make the change of variables $v_0 \mapsto (\rho_0, \theta_0, \varphi_0)$ and $v_1 \mapsto (\rho_1, \theta_1, \varphi_1)$ so that (3.31) becomes

$$\begin{aligned} \mathbf{E}_{(r_0, \Upsilon_0)}^\dagger[f(R_t, \Upsilon_t) \mathbf{1}_{\{J_1 \leq t < J_2\}}] &\geq C_1 \underline{\alpha} e^{-\bar{\alpha}t} \underline{\pi} \int_0^t ds \int_{v_{\min}}^{v_{\max}} d\rho_1 \int_0^\pi d\varphi_1 \int_0^{2\pi} d\theta_1 \int_{v_{\min}}^{v_{\max}} d\rho_0 \int_0^\pi d\varphi_0 \int_0^{2\pi} d\theta_0 \\ &\quad f(r_0 + \Theta_{\rho_0, \rho_1, \theta_1, \varphi_1}(s, \theta_0, \varphi_0), \tilde{\Theta}(\rho_1, \theta_1, \varphi_1)) \Delta(\rho_0, \theta_0, \varphi_0) \Delta(\rho_1, \theta_1, \varphi_1), \end{aligned} \quad (3.32)$$

where

$$\Theta_{\rho_0, \rho_1, \theta_1, \varphi_1}(s, \theta_0, \varphi_0) = \begin{bmatrix} s\rho_0 \sin \varphi_0 \cos \theta_0 + (t-s)\rho_1 \sin \varphi_1 \cos \theta_1 \\ s\rho_0 \sin \varphi_0 \sin \theta_0 + (t-s)\rho_1 \sin \varphi_1 \sin \theta_1 \\ s\rho_0 \cos \varphi_0 + (t-s)\rho_1 \cos \varphi_1 \end{bmatrix}, \quad (3.33)$$

represents the spatial variable $sv_0 + (t-s)v_1$ in polar coordinates,

$$\tilde{\Theta}(\rho_1, \theta_1, \varphi_1) = \begin{bmatrix} \rho_1 \sin \varphi_1 \cos \theta_1 \\ \rho_1 \sin \varphi_1 \sin \theta_1 \\ \rho_1 \cos \varphi_1 \end{bmatrix} \quad (3.34)$$

represents v_1 in polar coordinates,

$$\Delta(\rho, \theta, \varphi) = \rho^2 \sin \varphi, \quad (3.35)$$

is the determinant of the Jacobian matrix for the change of variables from Cartesian to polar coordinates, and C_1 is an unimportant normalising constant.

For fixed ρ_0, ρ_1, θ_1 and φ_1 , we first consider the part of (3.32) given by

$$(s, \theta_0, \varphi_0) \mapsto \int_0^t ds \int_0^\pi d\varphi_0 \int_0^{2\pi} d\theta_0 f(r_0 + \Theta_{\rho_0, \rho_1, \theta_1, \varphi_1}(s, \theta_0, \varphi_0), \tilde{\Theta}(\rho_1, \theta_1, \varphi_1)) \Delta(\rho_0, \theta_0, \varphi_0), \quad (3.36)$$

The Jacobian of $\Theta_{\rho_0, \rho_1, \theta_1, \varphi_1}$, as a function of (s, θ_0, φ_0) , is given by

$$\begin{bmatrix} \rho_0 \cos \theta_0 \sin \varphi_0 - \rho_1 \cos \theta_1 \sin \varphi_1 & -s \rho_0 \sin \theta_0 \sin \varphi_0 & s \rho_0 \cos \varphi_0 \cos \theta_0 \\ \rho_0 \sin \theta_0 \sin \varphi_0 - \rho_1 \sin \theta_1 \sin \varphi_1 & s \rho_0 \cos \theta_0 \sin \varphi_0 & s \rho_0 \cos \varphi_0 \sin \theta_0 \\ \rho_0 \cos \varphi_0 - \rho_1 \cos \varphi_1 & 0 & -s \rho_0 \sin \varphi_0 \end{bmatrix}.$$

whose determinant, $\det(D_{\rho_0, \rho_1, \theta_1, \varphi_1}(s, \theta_0, \varphi_0))$ satisfies

$$\frac{\Delta(\rho_0, \theta_0, \varphi_0)}{\det(D_{\rho_0, \rho_1, \theta_1, \varphi_1}(s, \theta_0, \varphi_0))} \geq \frac{1}{4s^2 v_{\max}^3} \geq \frac{1}{4t^2 v_{\max}^3}, \quad s \leq t.$$

We thus have the following lower bound for (3.36)

$$\begin{aligned} & \frac{1}{4t^2 v_{\max}^3} \int_0^t ds \int_0^\pi d\varphi_0 \int_0^{2\pi} d\theta_0 f(r_0 + \Theta_{\rho_0, \rho_1, \theta_1, \varphi_1}(s, \theta_0, \varphi_0), \tilde{\Theta}(\rho_1, \theta_1, \varphi_1)) \\ & \quad \times \det(D_{\rho_0, \rho_1, \theta_1, \varphi_1}(s, \theta_0, \varphi_0)). \end{aligned} \quad (3.37)$$

Making another change of variables $(s, \theta_0, \varphi_0) \mapsto r \in \mathbb{R}^3$ and using the fact that, regardless of the values of ρ_1, θ_1 and φ_1 , $\Theta_{\rho_0, \rho_1, \theta_1, \varphi_1}$ maps $(0, t) \times (0, \pi) \times (0, 2\pi)$ surjectively onto a set that contains $B(\rho_0 t)$, where $B(r)$ is the ball in \mathbb{R}^3 of radius r centred at the origin, (3.37), and hence (3.36), is bounded below by

$$\frac{1}{4t^2 v_{\max}^3} \int_{B(\rho_0 t)} f(r, \tilde{\Theta}(\rho_1, \theta_1, \varphi_1)) dr. \quad (3.38)$$

Substituting this equation back into (3.32) and changing $(\rho_1, \theta_1, \varphi_1)$ back to Cartesian coordinates, we have

$$\mathbf{E}_{(r_0, \Upsilon_0)}^\dagger [f(R_t, \Upsilon_t) \mathbf{1}_{\{J_1 \leq t < J_2\}}] \geq \frac{C_2 e^{-\bar{\alpha} t}}{t^2} \int_{v_{\min}}^{v_{\max}} d\rho_0 \int_{B(\rho_0 t)} dr \int_V dv_1 f(r, v_1), \quad (3.39)$$

where $C_2 = \underline{\alpha} \pi C_1 / (4v_{\max}^3)$.

Now suppose we fix an initial configuration $(r_0, v_0) \in D \times V$, with $tv_{\max} < \text{dist}(r_0, \partial D)$. By considering the event $\{J_2 \leq t < J_3\}$ and noting that the scattering kernel is bounded below

by $\underline{\pi}$, we may apply the Markov property together with (3.39) to the process at time J_1 before choosing the new velocity. Using the bounds on α and π as before, and recalling that Υ_0 is uniformly distributed, we have

$$\begin{aligned}
& \mathbf{E}_{(r_0, v_0)}^\dagger [f(R_t, \Upsilon_t) \mathbf{1}_{\{J_2 \leq t < J_3\}}] \\
& \geq \int_0^t ds \underline{\alpha} e^{-\bar{\alpha}s} \underline{\pi} \mathbf{E}_{(r_0 + sv_0, \Upsilon_0)}^\dagger [f(R_{t-s}, \Upsilon_{t-s}) \mathbf{1}_{\{J_1 \leq t-s < J_2\}}] \\
& \geq \int_0^t ds \underline{\alpha} e^{-\bar{\alpha}s} \underline{\pi} \frac{C_2 e^{-\bar{\alpha}(t-s)}}{(t-s)^2} \int_V dv_1 \int_{\mathbf{v}_{\min}}^{\mathbf{v}_{\max}} d\rho_0 \int_{\rho_0(t-s)B} dr f(r_0 + sv_0 + r, v_1) \\
& \geq \frac{C_3 e^{-\bar{\alpha}t}}{t^2} \int_0^t ds \int_V dv_1 \int_{\mathbf{v}_{\min}}^{\mathbf{v}_{\max}} d\rho_0 \int_{\rho_0(t-s)B} dr f(r_0 + sv_0 + r, v_1) \\
& = \frac{C_3 e^{-\bar{\alpha}t}}{t^2} \int_0^t ds \int_V dv_1 \int_{\mathbf{v}_{\min}}^{\mathbf{v}_{\max}} d\rho_0 \int_{r_0 + sv_0 + \rho_0(t-s)B} dy f(y, v_1), \tag{3.40}
\end{aligned}$$

where we have used the substitution $y = r_0 + sv_0 + r$ to obtain the final line and C_3 is another constant in $(0, \infty)$. Now note that for $s \leq \rho_0 t / (\rho_0 + \mathbf{v}_{\max})$ we have $r_0 + B(\rho_0 t - (\rho_0 + \mathbf{v}_{\max})s) \subset r_0 + sv_0 + B(\rho_0(t-s))$. Combining this with (3.40) and using Fubini, we have

$$\begin{aligned}
& \mathbf{E}_{(r_0, v_0)}^\dagger [f(R_t, \Upsilon_t) \mathbf{1}_{\{J_2 \leq t < J_3\}}] \\
& \geq \frac{C_3 e^{-\bar{\alpha}t}}{t^2} \int_V dv_1 \int_{\mathbf{v}_{\min}}^{\mathbf{v}_{\max}} d\rho_0 \int_{\mathbb{R}^3} \mathbf{1}_{\{0 \leq s \leq \frac{\rho_0}{\rho_0 + \mathbf{v}_{\max}} t\}} ds \int_{\mathbb{R}^3} dy \mathbf{1}_{\{|y-r_0| \leq \rho_0 t - (\rho_0 + \mathbf{v}_{\max})s\}} f(y, v_1) \\
& = \frac{C_3 e^{-\bar{\alpha}t}}{t^2} \int_V dv_1 \int_{\mathbf{v}_{\min}}^{\mathbf{v}_{\max}} d\rho_0 \int_{\mathbb{R}^3} ds \int_{\mathbb{R}^3} dy \mathbf{1}_{\{0 \leq s \leq \frac{\rho_0 t - |y-r_0|}{\rho_0 + \mathbf{v}_{\max}}\}} f(y, v_1) \\
& = \frac{C_3 e^{-\bar{\alpha}t}}{t^2} \int_V dv_1 \int_{\mathbf{v}_{\min}}^{\mathbf{v}_{\max}} d\rho_0 \int_{\mathbb{R}^3} dy \mathbf{1}_{\{|y-r_0| \leq \rho_0 t\}} \left(\frac{\rho_0 t - |y-r_0|}{\rho_0 + \mathbf{v}_{\max}} \right) f(y, v_1). \tag{3.41}
\end{aligned}$$

We finally compute the integral with respect to $\rho_0 \in (\mathbf{v}_{\min}, \mathbf{v}_{\max})$. In order to do so, we first note that since $\rho_0 < \mathbf{v}_{\max}$, the integrand in (3.41) is bounded below by

$$\frac{\rho_0 t - |y-r_0|}{2\mathbf{v}_{\max}}.$$

Absorbing $1/2\mathbf{v}_{\max}$ into the constant C_3 , applying Fubini and computing the ρ_0 integral yields

$$\begin{aligned}
\mathbf{E}_{(r_0, v_0)}^\dagger [f(R_t, \Upsilon_t)] & \geq \frac{C_3 e^{-\bar{\alpha}t}}{t^2} \int_V dv_1 \int_{\mathbb{R}^3} dy \left[\frac{t}{2} \left(\mathbf{v}_{\max}^2 - \left(\mathbf{v}_{\min} \vee \frac{|y-r|}{t} \right)^2 \right) \right. \\
& \quad \left. - |y-r| \left(\mathbf{v}_{\max} - \mathbf{v}_{\min} \vee \frac{|y-r|}{t} \right) \right] \mathbf{1}_{\{|y-r_0| \leq \mathbf{v}_{\max} t\}} f(y, v_1), \tag{3.42}
\end{aligned}$$

as required. \square

We now turn to the proof of (A1) under the assumptions of (B1) and (B2).

Proof of (A1). In this proof, we will follow a similar strategy to the one presented in [7, Section 4.2]. We therefore start by proving (A1) for initial configurations in $D_\varepsilon \times V$.

To this end, fix $(r, v) \in D_\varepsilon \times V$. From Lemma 6, there exists an $i \in \{1, \dots, \mathbf{n}\}$ such that $r \in B(r_i, \mathbf{v}_{\max}\varepsilon/32) \cap D_\varepsilon$. Then, for each $t \in [\varepsilon/2, \varepsilon)$, Lemma 7 yields

$$\begin{aligned} \mathbf{P}_{(r,v)}^\dagger(R_t \in dz, \Upsilon_t \in dw) &\geq \frac{Ce^{-\bar{\alpha}t}}{t^2} \left[\frac{t}{2} \left(\mathbf{v}_{\max}^2 - \left(\mathbf{v}_{\min} \vee \frac{|z-r|}{t} \right)^2 \right) \right. \\ &\quad \left. - |z-r| \left(\mathbf{v}_{\max} - \mathbf{v}_{\min} \vee \frac{|z-r|}{t} \right) \right] \mathbf{1}_{\{z \in B(r, \mathbf{v}_{\max}t)\}} dz dw. \end{aligned} \quad (3.43)$$

Now, if $j \in \{1, \dots, \mathbf{n}\}$ is such that $B(r_i, \mathbf{v}_{\max}\varepsilon/32) \cap B(r_j, \mathbf{v}_{\max}\varepsilon/32) \neq \emptyset$, the triangle inequality implies that $D_\varepsilon \cap (B(r_i, \mathbf{v}_{\max}\varepsilon/32) \cup B(r_j, \mathbf{v}_{\max}\varepsilon/32)) \subset B(r, \mathbf{v}_{\max}\varepsilon/8) \subset B(r, \mathbf{v}_{\max}t)$, with the latter inclusion following from the fact that $t \in [\varepsilon/2, \varepsilon)$.

Hence, for $z \in B(r_i, \mathbf{v}_{\max}\varepsilon/32) \cup B(r_j, \mathbf{v}_{\max}\varepsilon/32)$ and $t \in [\varepsilon/2, \varepsilon)$, the density on the right-hand side of (3.43) is bounded below by a constant $C_\varepsilon > 0$, which is independent of r, v, i and j . Hence,

$$\mathbf{P}_{(r,v)}^\dagger(R_t \in dz, \Upsilon_t \in dw) \geq C_\varepsilon \mathbf{1}_{\{z \in D_\varepsilon \cap (B(r_i, \varepsilon/32) \cup B(r_j, \varepsilon/32))\}} dz dw, \quad z \in D, w \in V. \quad (3.44)$$

Now let $t \geq (\mathbf{n}+1)\varepsilon/2$. By writing $t = k\varepsilon/2 + t'$, for some $k \geq \mathbf{n}$ and $t' \in [\varepsilon/2, \varepsilon)$. We will demonstrate that a repeated application of (3.44) will lead to the inequality

$$\mathbf{P}_{(r,v)}^\dagger(R_t \in dz, \Upsilon_t \in dw) \geq C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z \in D_\varepsilon\}} dz dw, \quad z \in D, w \in V, \quad (3.45)$$

for $(r, v) \in D_\varepsilon \times V$, where $c_\varepsilon > 0$ is another unimportant constant which depends only on ε and is defined in the following analysis.

To this end, we start by noting that, since $r \in D_\varepsilon$ and $v \in V$, there exists $i_0, i_1 \in \{1, \dots, \mathbf{n}\}$ such that $r \in B(r_{i_0}, \mathbf{v}_{\max}\varepsilon/32)$ and $B(r_{i_0}, \mathbf{v}_{\max}\varepsilon/32) \cap B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32) \cap D_\varepsilon \neq \emptyset$. Applying (3.44) at time t' (recall that we have identified $t = k\varepsilon/2 + t'$ for some $k \geq \mathbf{n}$) we obtain,

$$\begin{aligned} \mathbf{P}_{(r,v)}^\dagger(R_t \in dz, \Upsilon_t \in dw) &= \mathbf{P}_{(r,v)}^\dagger(R_{t'+k\varepsilon/2} \in dz, \Upsilon_{t'+k\varepsilon/2} \in dw) \\ &\geq \mathbf{E}_{(r,v)}^\dagger \left[\mathbf{1}_{\{R_{t'} \in B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32) \cap D_\varepsilon, \Upsilon_{t'} \in V\}} \mathbf{P}_{(R_{t'}, \Upsilon_{t'})}^\dagger(R_{k\varepsilon/2} \in dz, \Upsilon_{k\varepsilon/2} \in dw) \right] \\ &= \int_{B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32) \cap D_\varepsilon} \int_V \mathbf{P}_{(r',v')}^\dagger(R_{k\varepsilon/2} \in dz, \Upsilon_{k\varepsilon/2} \in dw) \mathbf{P}_{(r,v)}^\dagger(R_{t'} \in dr', \Upsilon_{t'} \in dv') \\ &\geq C_\varepsilon \int_{B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32) \cap D_\varepsilon} \int_V \mathbf{P}_{(r',v')}^\dagger(R_{k\varepsilon/2} \in dz, \Upsilon_{k\varepsilon/2} \in dw) \\ &\quad \times \mathbf{1}_{\{r' \in (B(r_{i_0}, \mathbf{v}_{\max}\varepsilon/32) \cup B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32)) \cap D_\varepsilon\}} dr' dv' \\ &= C_\varepsilon \int_{B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32) \cap D_\varepsilon} \int_V \mathbf{P}_{(r',v')}^\dagger(R_{k\varepsilon/2} \in dz, \Upsilon_{k\varepsilon/2} \in dw) dr' dv'. \end{aligned} \quad (3.46)$$

We now turn our attention to $\mathbf{P}_{(r',v')}^\dagger(R_{k\varepsilon/2} \in dz, \Upsilon_{k\varepsilon/2} \in dw)$, for $(r', v') \in (B(r_{i_1}, \mathbf{v}_{\max}\varepsilon/32) \cap D_\varepsilon) \times V$ and $k \geq \mathbf{n}$. Thanks to Lemma 6, for all $i_{k+1} \in \{1, \dots, \mathbf{n}\}$, there exist $i_2, \dots, i_k \in$

$\{1, \dots, \mathbf{n}\}$ such that $B(r_{i_j}, \varepsilon/32) \cap B(r_{i_{j+1}}, \varepsilon/32) \neq \emptyset$ for every $j \in \{1, \dots, k\}$. Note, here we see the importance of choosing $k \geq \mathbf{n}$, to ensure the validity of the previous statement.

Applying (3.44) and following the same steps that lead to (3.46), we obtain

$$\begin{aligned} & \mathbf{P}_{(r', v')}^\dagger(R_{k\varepsilon/2} \in dz, \Upsilon_{k\varepsilon/2} \in dw) \\ & \geq C_\varepsilon \int_{B(r_{i_2}, \varepsilon/32) \cap D_\varepsilon} \int_V \mathbf{P}_{(r'', v'')}^\dagger(R_{(k-1)\varepsilon/2} \in dz, \Upsilon_{(k-1)\varepsilon/2} \in dw) dr'' dv''. \end{aligned} \quad (3.47)$$

Iterating this step a further $k - 2$ times, we obtain

$$\begin{aligned} & \mathbf{P}_{(r', v')}^\dagger(R_{k\varepsilon/2} \in dz, \Upsilon_{k\varepsilon/2} \in dw) \\ & \geq C_\varepsilon c_\varepsilon^{k-2} \int_{B(r_{i_k}, v_{\max}\varepsilon/32) \cap D_\varepsilon} \int_V \mathbf{P}_{(r'', v'')}^\dagger(R_{\varepsilon/2} \in dz, \Upsilon_{\varepsilon/2} \in dw) dr'' dv'', \end{aligned} \quad (3.48)$$

where $c_\varepsilon = C_\varepsilon \text{Vol}(V) \min_{i=1, \dots, n} \text{Vol}(B(r_i, v_{\max}\varepsilon/32) \cap D_\varepsilon)$. Using this inequality to bound the right-hand side of (3.46) yields

$$\begin{aligned} & \mathbf{P}_{(r, v)}^\dagger(R_t \in dz, \Upsilon_t \in dw) \\ & \geq C_\varepsilon c_\varepsilon^{k-1} \int_{B(r_{i_k}, \varepsilon/32) \cap D_\varepsilon} \int_V \mathbf{P}_{(r', v')}^\dagger(R_{\varepsilon/2} \in dz, \Upsilon_{\varepsilon/2} \in dw) dr' dv'. \end{aligned} \quad (3.49)$$

We now apply (3.44) a final time at time $\varepsilon/2$ to obtain

$$\mathbf{P}_{(r, v)}^\dagger(R_t \in dz, \Upsilon_t \in dw) \geq C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z \in B(r_{i_{k+1}}, \varepsilon/2) \cap D_\varepsilon\}} dz dw. \quad (3.50)$$

Since this inequality holds for every $i_{k+1} \in \{1, \dots, \mathbf{n}\}$, it also follows that

$$\begin{aligned} \mathbf{P}_{(r, v)}^\dagger(R_t \in dz, \Upsilon_t \in dw) & \geq C_\varepsilon c_\varepsilon^k \sup_{i_{k+1} \in \{1, \dots, \mathbf{n}\}} \mathbf{1}_{\{z \in B(r_{i_{k+1}}, \varepsilon/2) \cap D_\varepsilon\}} dz dw \\ & \geq C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z \in D_\varepsilon\}} dz dw, \end{aligned}$$

where the final line follows from Lemma 6 since $k + 1 > \mathbf{n}$. This is the lower bound claimed in (3.45).

Finally, noting that for any two events A, B , $\Pr(A|B) = \Pr(A \cap B) / \Pr(B) \geq \Pr(A \cap B)$, we have that for initial conditions $(r, v) \in D_\varepsilon \times V$, any $t_0 \geq (\mathbf{n} + 1)\varepsilon/2$ and ν equal to Lebesgue measure on $D_\varepsilon \times V$, there exists a constant $c_1 \in (0, \infty)$ such that

$$\mathbf{P}_{(r, v)}((R_{t_0}, \Upsilon_{t_0}) \in \cdot | t_0 < \mathbf{k}) \geq c_1 \nu(\cdot),$$

as required by (A1).

We now prove (A1) for initial conditions in $(D \setminus D_\varepsilon) \times V$. Once again, we recall that assumptions (B1) and (B2) are in force.

Choose $r \in D \setminus D_\varepsilon$, $v \in V$ and define the (deterministic) time

$$\kappa_{r,v}^{D \setminus D_\varepsilon} := \inf\{t > 0 : r + tv \notin \partial D \setminus D_\varepsilon\},$$

which is the time it would take a neutron released at r with velocity v to hit the boundary of $D \setminus D_\varepsilon$ if no scatter or fission took place. Note in particular that $\kappa_{r,v}^{D \setminus D_\varepsilon}$ is not a random time but entirely deterministic. We first consider the case $r + \kappa_{r,v}^{D \setminus D_\varepsilon} v \in \partial D_\varepsilon$

$$\mathbf{P}_{(r,v)}^\dagger(R_{\kappa_{r,v}^{D \setminus D_\varepsilon}} \in \partial D_\varepsilon) \geq e^{-\bar{\alpha}\kappa_{r,v}^{D \setminus D_\varepsilon}} \geq e^{-\bar{\alpha}\text{diam}(D)/v_{\min}}. \quad (3.51)$$

Combining this with (3.45) and the Markov property, for all $t \geq (\mathbf{n} + 1)\varepsilon/2$

$$\begin{aligned} & \mathbb{P}_{(r,v)}(R_{\kappa_{r,v}^{D \setminus D_\varepsilon} + t} \in dz, \Upsilon_{\kappa_{r,v}^{D \setminus D_\varepsilon} + t} \in dw | \kappa_{r,v}^{D \setminus D_\varepsilon} + t < \mathbf{k}) \\ & \geq \mathbf{P}_{(r,v)}^\dagger(R_{\kappa_{r,v}^{D \setminus D_\varepsilon} + t} \in dz, \Upsilon_{\kappa_{r,v}^{D \setminus D_\varepsilon} + t} \in dw) \\ & \geq e^{-\bar{\alpha}\text{diam}(D)/v_{\min}} C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z \in D_\varepsilon\}} dz dw, \end{aligned} \quad (3.52)$$

where $k \geq \mathbf{n}$ is such that $t = k\varepsilon/2 + t'$ for some $t' \in [\varepsilon/2, \varepsilon)$.

On the other hand, suppose $r + \kappa_{r,v}^{D \setminus D_\varepsilon} v \in \partial D$. Then, recalling the assumptions (B1) and (B2) it follows that $\{J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon), \Upsilon_{J_1} \in K_{r+vJ_1}, J_2 > t_\varepsilon\} \subset \{R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathbf{k}\}$. Heuristically speaking, this is because if the first jump occurs before time $\kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon)$, then the process hasn't hit the boundary and there are still (at least) s_ε units of time left until t_ε . By then choosing the new velocity, Υ_{J_1} , from K_{r+vJ_1} , thanks to the assumption (B1) and the remarks around (3.29), this implies that the process will remain in $D \setminus D_\varepsilon$ for s_ε units of time, at some point in time after which, it will move into D_ε , providing the process doesn't jump again before entering D_ε . Combining this with the usual bounds on α , and recalling from (B2) that $\text{Vol}(K_r) > \gamma > 0$ for all $r \in D \setminus D_\varepsilon$ and $v \in V$, we have

$$\begin{aligned} \mathbf{P}_{(r,v)}(R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathbf{k}) & \geq \mathbf{P}_{(r,v)}^\dagger(J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon), \Upsilon_{J_1} \in K_{r+vJ_1}, J_2 > t_\varepsilon) \\ & \geq \underline{\pi}\gamma e^{-\bar{\alpha}t_\varepsilon} \mathbf{P}_{(r,v)}^\dagger(J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon)). \end{aligned} \quad (3.53)$$

Along with (3.45), this implies that, for all $r \in D \setminus D_\varepsilon$, $v \in V$ and $t \geq (\mathbf{n} + 1)\varepsilon/2$ such that

$$t + t_\varepsilon \geq \kappa_{r,v}^{D \setminus D_\varepsilon}$$

$$\begin{aligned}
& \mathbf{P}_{(r,v)}(R_{t+t_\varepsilon} \in dz, \Upsilon_{t+t_\varepsilon} \in dw | t + t_\varepsilon < \mathbf{k}) \\
& \geq \frac{\mathbf{P}_{(r,v)}(R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathbf{k}, R_{t+t_\varepsilon} \in dz, \Upsilon_{t+t_\varepsilon} \in dw)}{\mathbf{P}_{(r,v)}(t + t_\varepsilon < \mathbf{k})} \\
& = \frac{\mathbf{P}_{(r,v)}^\dagger(R_{t+t_\varepsilon} \in dz; \Upsilon_{t+t_\varepsilon} \in dw | R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathbf{k}) \mathbf{P}_{(r,v)}(R_{t_\varepsilon} \in D_\varepsilon, t_\varepsilon < \mathbf{k})}{\mathbf{P}_{(r,v)}(t + t_\varepsilon < \mathbf{k})} \\
& \geq \inf_{r \in D_\varepsilon, v \in V} \mathbf{P}_{(r,v)}^\dagger(R_t \in dz; \Upsilon_t \in dw) \\
& \quad \times \frac{\mathbf{P}_{(r,v)}^\dagger(J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,v)}(t + t_\varepsilon < \mathbf{k})} \pi \gamma e^{-\bar{\alpha} t_\varepsilon} c_\varepsilon^k \mathbf{1}_{\{z \in D_\varepsilon\}} dz dw \\
& \geq \frac{\mathbf{P}_{(r,v)}^\dagger(J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,v)}(t + t_\varepsilon < \mathbf{k})} \pi \gamma e^{-\bar{\alpha} t_\varepsilon} C_\varepsilon c_\varepsilon^k \mathbf{1}_{\{z \in D_\varepsilon\}} dz dw. \tag{3.54}
\end{aligned}$$

Now, since we are considering the case $r + \kappa_{r,v}^{D \setminus D_\varepsilon} v \in \partial D$ and $t + t_\varepsilon \geq \kappa_{r,v}^{D \setminus D_\varepsilon}$, it follows that $\{t + t_\varepsilon < \mathbf{k}\} \subset \{J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon}\}$. Then,

$$\begin{aligned}
\frac{\mathbf{P}_{(r,v)}^\dagger(J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,v)}(t + t_\varepsilon < \mathbf{k})} & \geq \frac{\mathbf{P}_{(r,v)}^\dagger(J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon))}{\mathbf{P}_{(r,v)}(J_1 < \kappa_{r,v}^{D \setminus D_\varepsilon})} \\
& \geq \frac{1 - e^{-\underline{\alpha}(\kappa_{r,v}^{D \setminus D_\varepsilon} \wedge (t_\varepsilon - s_\varepsilon))}}{1 - e^{-\bar{\alpha} \kappa_{r,v}^{D \setminus D_\varepsilon}}}, \tag{3.55}
\end{aligned}$$

with the bound on the right-hand side above being itself bounded below by a constant that does not depend on (r, v) . Substituting this back into (3.54), this proves (A1) with ν taken as Lebesgue measure on $D_\varepsilon \times V$ as before, t_0 can be sufficiently taken as $(\mathbf{n} + 1)\varepsilon/2 + \text{diam}(D)/\mathbf{v}_{\min}$ and we may start with any initial configurations in $D \setminus D_\varepsilon \times V$. \square

In order to prove (A2) we require the following lemma, the proof of which will be given after that of (A2).

Lemma 8. *For all $r \in D$ and $v \in V$, recalling that J_k denotes the k^{th} jump time of the process (R, Υ) , we have*

$$\mathbf{P}_{(r,v)}^\dagger(J_7 < \mathbf{k}, R_{J_7} \in dz) \leq C \mathbf{1}_{\{z \in D\}} dz, \tag{3.56}$$

for some constant $C > 0$, and

$$\mathbf{P}_\nu^\dagger(J_1 < \mathbf{k}, R_{J_1} \in dz) \geq c \mathbf{1}_{\{z \in D\}} dz, \tag{3.57}$$

for another constant $c > 0$, where ν , from the proof of (A1), is Lebesgue measure on $D_\varepsilon \times V$.

Proof of (A2). Again, we follow the proof given by the authors in [7]. Let $t \geq 7\text{diam}(D)/\mathbf{v}_{\min}$ and note that on the event $\{\mathbf{k} > t\}$, we have $J_7 \leq 7\text{diam}(D)/\mathbf{v}_{\min}$ almost surely. This inequality

along with the strong Markov property imply that,

$$\begin{aligned} \mathbf{P}_{(r,v)}(t < \mathbf{k}) &\leq \mathbf{E}_{(r,v)}^\dagger \left[\mathbf{1}_{\{J_7 < t\}} \mathbf{P}_{(R_{J_7}, \Upsilon_{J_7})}(t - s < \mathbf{k})_{s=J_7} \right] \\ &\leq \mathbf{E}_{(r,v)}^\dagger \left[\mathbf{P}_{(R_{J_7}, \Upsilon_{J_7})} \left(t - \frac{7\text{diam}(D)}{\mathbf{v}_{\min}} < \mathbf{k} \right) \right]. \end{aligned} \quad (3.58)$$

Since π is uniformly bounded above, conditional on $\{J_7 < \infty, R_{J_7} \in dz\}$, the density of Υ_{J_7} is bounded above by $\bar{\pi}$ multiplied by Lebesgue measure on V . Combining this with (3.56) and (3.58), we obtain

$$\mathbf{P}_{(r,v)}(t < \mathbf{k}) \leq C' \int_D \int_V \mathbf{P}_{(z,w)} \left(t - \frac{7\text{diam}(D)}{\mathbf{v}_{\min}} < \mathbf{k} \right) dw dz, \quad (3.59)$$

for some $C' \in (0, \infty)$. Similarly, for $t \geq \text{diam}(D)/\mathbf{v}_{\min}$, equation (3.57), the fact that the inclusion $\{t < \mathbf{k}\} \subset \{J_1 \leq \text{diam}(D)/\mathbf{v}_{\min}\}$, the strong Markov property and the fact that π is uniformly bounded below entail that,

$$\begin{aligned} \mathbf{P}_\nu(t < \mathbf{k}) &= \mathbf{E}_\nu^\dagger \left[\mathbf{1}_{\{J_1 \leq \mathbf{k}\}} \mathbf{P}_{(R_{J_1}, \Upsilon_{J_1})}(t - s < \mathbf{k})_{s=J_1} \right] \\ &\geq \mathbf{E}_\nu^\dagger \left[\mathbf{1}_{\{J_1 \leq \mathbf{k}\}} \mathbf{P}_{(R_{J_1}, \Upsilon_{J_1})}(t < \mathbf{k}) \right] \\ &\geq c' \int_D \int_V \mathbf{P}_{(z,w)}(t < \mathbf{k}) dw dz, \end{aligned}$$

for some $c' \in (0, \infty)$, where ν is Lebesgue measure on $D_\varepsilon \times V$. Putting (3.58) and (3.59) together, for all $t \geq 8\text{diam}(D)/\mathbf{v}_{\min}$, we have

$$\mathbf{P}_{(r,v)}(t < \mathbf{k}) \leq \frac{C'}{c'} \mathbf{P}_\nu \left(t - \frac{7\text{diam}(D)}{\mathbf{v}_{\min}} < \mathbf{k} \right). \quad (3.60)$$

Now, recalling t_0 and ν from the proof of (A1), it follows from (A1) that

$$\mathbf{P}_\nu^\dagger((R_{t_0}, \Upsilon_{t_0}) \in \cdot) \geq c_1 \mathbf{P}_\nu(t_0 < \mathbf{k}) \nu(\cdot). \quad (3.61)$$

The event $\{t < \mathbf{k}\}$ occurs if the particle has either been killed on the boundary of D or if it has been absorbed by fissile material, which occurs at rate $\bar{\beta} - \beta$. Since t_0 and ν are fixed, and $\bar{\beta} - \beta \leq \bar{\beta} + 1 < \infty$ by assumption, $\mathbf{P}_\nu(t_0 < \mathbf{k}) \geq K$ for some constant $K > 0$. Thus, keeping $t \geq 8\text{diam}(D)/\mathbf{v}_{\min}$, using (3.61)

$$\begin{aligned} \mathbf{P}_\nu \left(t - \frac{7\text{diam}(D)}{\mathbf{v}_{\min}} + t_0 < \mathbf{k} \right) &= \mathbf{E}_\nu \left[\mathbf{1}_{\{t_0 < \mathbf{k}\}} \mathbf{P}_{(R_{t_0}, \Upsilon_{t_0})}^\dagger \left(t - \frac{7\text{diam}(D)}{\mathbf{v}_{\min}} < \mathbf{k} \right) \right] \\ &\geq \tilde{c}_1 \mathbf{P}_\nu \left(t - \frac{7\text{diam}(D)}{\mathbf{v}_{\min}} < \mathbf{k} \right), \end{aligned} \quad (3.62)$$

where $\tilde{c}_1 = Kc_1$.

Now define $N = \lceil 7\text{diam}(D)/(\mathbf{v}_{\min}t_0) \rceil$. Then, for any $t > 0$, $t - 7\text{diam}(D)/\mathbf{v}_{\min} + Nt_0 \geq t$

so that, trivially,

$$\mathbf{P}_\nu(t < \mathbf{k}) \geq \mathbf{P}_\nu \left(t - \frac{7\text{diam}(D)}{\mathbf{v}_{\min}} + Nt_0 < \mathbf{k} \right). \quad (3.63)$$

Applying (3.62) N times implies that

$$\mathbf{P}_\nu(t < \mathbf{k}) \geq \tilde{c}_1^N \mathbf{P}_\nu \left(t - \frac{4\text{diam}(D)}{\mathbf{v}_{\min}} < \mathbf{k} \right). \quad (3.64)$$

Combining this with (3.60) completes the proof of (A2). \square

Proof of Lemma 8. Let us first prove (3.56). Again, following the proof given in [7], we couple the neutron transport random walk in D with one on the whole of \mathbb{R}^3 . Denote by $(\hat{R}_t, \hat{\Upsilon}_t)$ the neutron random walk in $\hat{D} = \mathbb{R}^3$, coupled with (R, Υ) such that $\hat{R}_t = R_t$ and $\hat{\Upsilon}_t = \Upsilon_t$ for all $t < \mathbf{k}$ and $(R_0, \Upsilon_0) = (\hat{R}_0, \hat{\Upsilon}_0) = (r, v)$, for $r \in D$, $v \in V$. Denote by $\hat{J}_1 < \hat{J}_2 < \dots$ the jump times of $\hat{\Upsilon}_t$. Then for each $k \geq 1$ such that $J_k < \mathbf{k}$, we have $\hat{J}_k = J_k$. Due to the inequality

$$\mathbf{E}_{(r,v)}^\dagger[f(R_{J_7}); J_7 < \mathbf{k}] \leq \mathbf{E}_{(r,v)}[f(\hat{R}_{\hat{J}_7})], \quad r \in D, v \in V, \quad (3.65)$$

we will consider the distribution of $\hat{R}_{\hat{J}_i}$ for $i \geq 2$. We first look at the case when $i = 2$. For $(r, v) \in D \times V$ and non-negative, bounded, measurable functions f ,

$$\begin{aligned} \mathbf{E}_{(r,v)}[f(\hat{R}_{\hat{J}_2})] &= \mathbf{E}_{(r,v)}[f(r + v\hat{J}_1 + \hat{\Upsilon}_{\hat{J}_1}(\hat{J}_2 - \hat{J}_1))] \\ &\leq \bar{\alpha}^2 \bar{\pi} \int_0^\infty dj_1 \int_V dv_1 \int_0^\infty dj_2 e^{-\alpha(j_1+j_2)} f(r + vj_1 + v_1j_2) \end{aligned} \quad (3.66)$$

For j_1 fixed, we consider the integrals over v_1 and j_2 in (3.66). Making the change of variables $v_1 \mapsto (\rho, \varphi, \theta)$, we have

$$\begin{aligned} &\int_V dv_1 \int_0^\infty dj_2 e^{-\alpha j_2} f(r + vj_1 + v_1j_2) \\ &\leq \int_{\mathbf{v}_{\min}}^1 d\rho \int_0^{2\pi} d\theta \int_0^\pi d\varphi \int_0^\infty dj_2 e^{-\alpha j_2} f\left(r + vj_1 + \tilde{\Theta}(\rho j_2, \theta, \varphi)\right) \rho^2 \sin \varphi, \end{aligned} \quad (3.67)$$

where $\tilde{\Theta}$ was defined in (3.34). Now making the substitution $u = \rho j_2$ in (3.67),

$$\begin{aligned} &\int_V dv_1 \int_0^\infty dj_2 e^{-\alpha j_2} f(r + vj_1 + v_1j_2) \\ &\leq \int_{\mathbf{v}_{\min}}^{\mathbf{v}_{\max}} d\rho \int_0^{2\pi} d\theta \int_0^\pi d\varphi \int_0^\infty du e^{-\alpha u/\rho} f\left(r + vj_1 + \tilde{\Theta}(u, \theta, \varphi)\right) \rho \sin \varphi \\ &\leq C \int_0^{2\pi} d\theta \int_0^\pi d\varphi \int_0^\infty du e^{-\alpha u/\mathbf{v}_{\max}} f\left(r + vj_1 + \tilde{\Theta}(u, \theta, \varphi)\right) \sin \varphi, \end{aligned} \quad (3.68)$$

where $C = \mathbf{v}_{\max}(\mathbf{v}_{\max} - \mathbf{v}_{\min})$. Making a final change of variables $(u, \theta, \varphi) \mapsto x \in \mathbb{R}^3$, we have

$$\int_V dv_1 \int_0^\infty dj_2 e^{-\alpha j_2} f(r + vj_1 + v_1j_2) \leq C \int_{\mathbb{R}^3} dx f(r + vj_1 + x) \frac{e^{-\alpha|x|/\mathbf{v}_{\max}}}{|x|^2}. \quad (3.69)$$

Substituting this back into (3.66) yields

$$\mathbb{E}_{(r,v)}[f(\hat{R}_{j_2})] \leq \bar{\alpha}K \int_0^\infty dj_1 e^{-\alpha j_1} \int_{\mathbb{R}^3} dx f(r + vj_1 + x) \frac{e^{-\alpha|x|/v_{\max}}}{|x|^2}, \quad (3.70)$$

where $K = \bar{\alpha}\pi C$. Iterating this process over the next five jumps of the process gives

$$\mathbb{E}_{(r,v)}[f(\hat{R}_{j_7})] \leq \bar{\alpha}K^6 \int_0^\infty dj_1 e^{-\alpha j_1} \int_{\mathbb{R}^3} dx_1 \cdots \int_{\mathbb{R}^3} dx_6 f(r + vj_1 + x_1 + \cdots + x_6) g(x_1) \cdots g(x_6) \quad (3.71)$$

where $g(x) = e^{-\alpha|x|/v_{\max}}/|x|^2$, $x \in \mathbb{R}^3$. Now, $g \in L^p(\mathbb{R}^3)$ for each $p < 3/2$ so that, in particular, $g \in L^{6/5}(\mathbb{R}^3)$. Hence, repeatedly applying Young's convolution inequality implies that the six-fold convolution $*^6 g \in L^\infty(\mathbb{R}^3)$. (The reader will note that this is the fundamental reason we have focused our calculations around the 7th jump time J_7 , rather than it being an arbitrary choice.) Making the substitution $x = x_1 + \cdots + x_6$,

$$\mathbf{E}_{(r,v)}^\dagger[f(\hat{R}_{j_7})] \leq \bar{\alpha}K^6 \|*^6 g\|_\infty \int_0^\infty dj_1 e^{-\alpha j_1} \int_{\mathbb{R}^3} dx_1 \cdots \int_{\mathbb{R}^3} dx_6 f(r + vj_1 + x). \quad (3.72)$$

Finally, setting $z = r + vj_1 + x$ yields

$$\mathbf{E}_{(r,v)}^\dagger[f(R_{J_7}); J_7 < \mathbf{k}] \leq \mathbb{E}_{(r,v)}[f(\hat{R}_{j_7})] \leq C' \int_{\mathbb{R}^3} f(z) dz, \quad (3.73)$$

where $C' = \bar{\alpha}K^6 \|g * \cdots * g\|_\infty$, which completes the proof of (3.56).

We now prove (3.57). For $r, r' \in \mathbb{R}^3$, let $[r, r']$ denote the line segment between r and r' . For all $f \in \mathcal{B}(\mathbb{R}^3)$, recalling the definition of ν from the proof of (A1) and using the usual bounds on α ,

$$\mathbf{E}_\nu[f(R_{J_1}); J_1 < \mathbf{k}] \geq \int_{D_\varepsilon} \frac{dr}{\text{Vol}(D_\varepsilon)} \int_V \frac{dv}{\text{Vol}(V)} \int_0^\infty ds \mathbf{1}_{\{[r, r+sv] \subset D\}} \underline{\alpha} e^{-\bar{\alpha}s} f(r + sv), \quad (3.74)$$

where $\text{Vol}(D_\varepsilon) = \int_{D_\varepsilon} dr$ and $\text{Vol}(V) = \int_V dv$. Following a similar method to those employed in the proof of Lemma 7 and (3.56) and changing first to polar coordinates via $v \mapsto (\rho, \theta, \varphi)$, followed by the substitution $u = s\rho$, and finally changing back to Cartesian coordinates via $(u, \theta, \varphi) \mapsto x$, the right-hand side of (3.74) is bounded below by

$$C \int_{D_\varepsilon} dr \int_{\mathbb{R}^3} dx \mathbf{1}_{\{[r, r+x] \subset D\}} \frac{\underline{\alpha} e^{-\bar{\alpha}|x|/v_{\min}}}{|x|^2} f(r + x), \quad (3.75)$$

where $C > 0$ is a constant. Making a final substitution of $x = z - r$, yields

$$\begin{aligned} \mathbf{E}_\nu[f(R_{J_1}); J_1 < \mathbf{k}] &\geq C \int_D dz \mathbf{1}_{[r, z] \subset D} \frac{\underline{\alpha} e^{-\bar{\alpha}|z-r|/v_{\min}}}{|z-r|^2} f(z) \\ &\geq C \frac{v_{\min}^2 \underline{\alpha} e^{-\bar{\alpha} \text{diam}(D)/v_{\min}^2}}{(\text{diam}(D))^2} \int_D dz \mathbf{1}_{[r, z] \subset D} f(z). \end{aligned} \quad (3.76)$$

For all $z \in D \setminus D_\varepsilon$, (B1) and the discussion thereafter now imply that

$$\int_{D_\varepsilon} \mathbf{1}_{\{[r,z] \subset D\}} dr \geq \text{Vol}(L_z) \geq \frac{\gamma}{2}(t_\varepsilon^2 - s_\varepsilon^2), \quad (3.77)$$

where s_ε and t_ε are defined in (B2), and L_z is defined in (3.29). On the other hand, for all $z \in D_\varepsilon$,

$$\int_{D_\varepsilon} \mathbf{1}_{\{[r,z] \subset D\}} dr \geq \text{Vol}(D_\varepsilon \cap B(r, \varepsilon)). \quad (3.78)$$

Since the map $z \mapsto \text{Vol}(D_\varepsilon \cap B(z, \varepsilon))$ is continuous and positive on the compact set \bar{D}_ε , the latter equation is uniformly bounded below by a strictly positive constant. It then follows that for every $z \in D$, the integral $\int_{D_\varepsilon} dr \mathbf{1}_{\{[r,z] \subset D\}}$ is bounded below by a positive constant. Using this to bound the right-hand side of (3.76) yields the result. \square

We thus have proved that the conclusions of Theorem 10 are valid under our assumptions. In order to conclude that Theorem 7 holds true, it remains to prove that φ is uniformly bounded away from 0 on each compactly embedded subset of $D \times V$ and the existence of a positive bounded density for the left eigenmeasure η .

Lemma 9. *The right eigenfunction φ is uniformly bounded away from 0 on each compactly embedded subset of $D \times V$ and the probability measure η admits a positive density with respect to the Lebesgue measure on $D \times V$, which corresponds to the quantity $\tilde{\varphi}$ and which is uniformly bounded from above and a.e. uniformly bounded from below on each compactly embedded subset of $D \times V$.*

Proof. For all $\varepsilon > 0$, we deduce from the eigenfunction property of φ (cf. Theorem 10) and from (3.45) that there exist a time $t_\varepsilon > 0$ and a constant $\tilde{C}_\varepsilon > 0$ such that

$$\varphi(r, v) = e^{-\lambda_c t_\varepsilon} \mathbf{P}_{t_\varepsilon}[\varphi](r, v) \geq e^{-\lambda t_\varepsilon} \tilde{C}_\varepsilon \int_{D_\varepsilon \times V} \varphi(z, w) dz dw > 0,$$

for all $(r, v) \in D_\varepsilon \times V$. It follows that φ is uniformly bounded away from 0 on each compactly embedded domain of $D \times V$.

Using the same notations as in the proof of Lemma 8, we consider the neutron transport random walk (\hat{R}_t, \hat{Y}_t) in $\hat{D} = \mathbb{R}^3$, coupled with (R, Υ) such that $\hat{R}_t = R_t$ and $\hat{Y}_t = \Upsilon_t$ for all $t < k$. We also denote by $\hat{J}_1 < \hat{J}_2 < \dots$ the jump times of $(\hat{Y}_t)_{t \geq 0}$. Let $T \geq 0$ be a random time independent of (\hat{R}, \hat{Y}) with uniform law on $[\underline{T}, \bar{T}]$, where $\underline{T} < \bar{T}$ are fixed and $\underline{T} \geq 7 \text{diam}(D)/v_{\min}$. We first prove that the law of (\hat{R}_T, \hat{Y}_T) after the 7th jump admits a uniformly bounded density with respect to the Lebesgue measure. We conclude by using the coupling with (R, Υ) and the quasi-stationary property of η in (3.26).

For all $k \geq 7$ and for any positive, bounded and measurable function f vanishing outside

of $D \times V$, we have

$$\begin{aligned}
& \mathbf{E}[f(\hat{R}_T, \hat{Y}_T) \mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{Y}_0, T] \\
&= \mathbf{E}[f(\hat{R}_0 + \hat{J}_1 \hat{Y}_0 + \cdots + \hat{J}_k \hat{Y}_{k-1} + (T - \hat{J}_1 - \cdots - \hat{J}_k) \hat{Y}_k, \hat{Y}_k) \mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{Y}_0, T] \\
&= \int_0^T ds_1 \alpha(\hat{R}_0 + v_0 s_1, v_0) e^{-\int_0^{s_1} \alpha(\hat{R}_0 + v_0 u, v_0) du} \\
&\quad \times \int_V dv_1 \pi(r_0 + v_0 s_1, v_0, v_1) \times \int_0^{T-s_1} ds_2 \alpha(\hat{R}_0 + v_0 s_1 + v_1 s_2, v_1) e^{-\int_0^{s_2} \alpha(\hat{R}_0 + v_0 s_1 + v_1 u, v_1) du} \\
&\quad \times \cdots \\
&\quad \times \int_V dv_{k-1} \pi(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-2} s_{k-1}, v_{k-2}, v_{k-1}) \\
&\quad \times \int_0^{T-s_1-\cdots-s_{k-1}} ds_k \alpha(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-1} s_k, v_{k-1}) \\
&\quad \quad \quad e^{-\int_0^{s_k} \alpha(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-2} s_{k-1} + v_{k-1} u, v_{k-1}) du} \\
&\quad \times \int_V dv_k \pi(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-1} s_k, v_{k-1}, v_k) \\
&\quad \times e^{-\int_0^{T-s_1-\cdots-s_k} \alpha(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-1} s_k + v_k u, v_k) du} \\
&\quad \quad \quad \times f(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-1} s_k + v_k(t - s_1 - \cdots - s_k), v_k).
\end{aligned}$$

Henceforth

$$\begin{aligned}
& \mathbf{E}[f(\hat{R}_T, \hat{Y}_T) \mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{Y}_0, T] \\
&\leq \bar{\alpha}^k \bar{\pi}^k e^{-T\bar{\alpha}} \int_0^T ds_1 \int_V dv_1 \cdots \int_0^{T-s_1-\cdots-s_{k-1}} ds_k \int_V dv_k \\
&\quad \times f(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-1} s_k + v_k(T - s_1 - \cdots - s_k), v_k).
\end{aligned}$$

Taking the expectation with respect to T , we obtain

$$\begin{aligned}
& \mathbf{E}[f(\hat{R}_T, \hat{Y}_T) \mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{Y}_0] \\
&\leq \frac{\bar{\alpha}^k \bar{\pi}^k}{\bar{T}} \int_0^{\bar{T}} dt \int_0^t ds_1 \int_V dv_1 \cdots \int_0^{t-s_1-\cdots-s_{k-1}} ds_k \int_V dv_k \\
&\quad \times f(\hat{R}_0 + v_0 s_1 + \cdots + v_{k-1} s_k + v_k(t - s_1 - \cdots - s_k), v_k).
\end{aligned}$$

Using the change of variable $(u_1, \dots, u_k, u_{k+1}) = (s_1, \dots, s_k, t - s_1 - \cdots - s_k)$ yields

$$\begin{aligned}
& \mathbf{E}[f(\hat{R}_T, \hat{Y}_T) \mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{Y}_0] \\
&\leq \frac{\bar{\alpha}^k \bar{\pi}^k}{\bar{T}} \int_{[0, \bar{T}]^{k+1}} du \mathbf{1}_{0 \leq u_1 + \cdots + u_{k+1} \leq \bar{T}} \int_{V^k} dv \\
&\quad \times f(\hat{R}_0 + v_0 u_1 + \cdots + v_{k-1} u_k + v_k u_{k+1}, v_k).
\end{aligned}$$

The same approach as in Lemma 8 shows that there exists a constant $C > 0$ (which does not depend on \hat{R}_0 nor on $\hat{\Upsilon}_0$) such that, for all measurable function $g : \mathbb{R}^3 \rightarrow [0, +\infty)$,

$$\int_{[0, \bar{T}]^7} du \int_{V^6} dv g(\hat{R}_0 + \hat{\Upsilon}_0 u_1 + \cdots + v_6 u_7) \leq C \int_{\mathbb{R}^d} dx g(x).$$

Hence,

$$\begin{aligned} & \mathbf{E}[f(\hat{R}_T, \hat{\Upsilon}_T) \mathbf{1}_{\{\hat{J}_k \leq T < \hat{J}_{k+1}\}} \mid \hat{R}_0, \hat{\Upsilon}_0] \\ & \leq \frac{C \bar{\alpha}^k \bar{\pi}^k}{\bar{T}} \int_{[0, \bar{T}]^{k+1-7}} du \mathbf{1}_{0 \leq u_8 + \cdots + u_{k+1} \leq \bar{T}} \int_{V^{k-6}} dv \\ & \quad \times \int_{\mathbb{R}^3} dx f(x + v_7 u_8 + \cdots + v_k u_{k+1}, v_k) \\ & = \frac{C \bar{\alpha}^k \bar{\pi}^k}{\bar{T}} \int_{[0, \bar{T}]^{k+1-7}} du \mathbf{1}_{0 \leq u_8 + \cdots + u_{k+1} \leq \bar{T}} \int_{V^{k-6}} dv \int_{\mathbb{R}^3} dy f(y, v_k) \\ & = C \bar{\alpha}^k \bar{\pi}^k \text{Vol}(V)^{k-8} \frac{\bar{T}^{k+1-8}}{(k+1-7)!} \int_D dy \int_V dv_k f(y, v_k) \end{aligned}$$

where we used the change of variable $y = x + v_7 u_8 + \cdots + v_k u_{k+1}$ and the fact that f vanishes outside $D \times V$. Summing over $k \geq 7$, we deduce that there exists a constant $C' > 0$ (which only depends on $C, \bar{\alpha}, \bar{\pi}$ and \bar{T}) such that

$$\mathbf{E}[f(\hat{R}_T, \hat{\Upsilon}_T) \mathbf{1}_{\{\hat{J}_7 \leq T\}} \mid \hat{R}_0, \hat{\Upsilon}_0] \leq C' \int_D dy \int_V dv f(y, v).$$

Similarly as in the proof of (A2), we chose $\underline{T} \geq 7 \text{diam}(D)/v_{\min}$, so that, on the event $\{\mathbf{k} > T\}$, we have $J_7 \leq 7 \text{diam}(D)/v_{\min} \leq T$ almost surely. Hence, we obtain that, for any $(r_0, v_0) \in D \times V$,

$$\begin{aligned} \mathbf{E}_{(r_0, v_0)}^\dagger[f(R_T, \Upsilon_T); T < \mathbf{k}] &= \mathbf{E}_{(r_0, v_0)}^\dagger[f(R_T, \Upsilon_T); T < \mathbf{k}, J_7 \leq T] \\ &\leq \mathbf{E}_{(r_0, v_0)}^\dagger[f(\hat{R}_T, \hat{\Upsilon}_T); \hat{J}_7 \leq T] \\ &\leq C' \int_D dy \int_V dv f(y, v). \end{aligned}$$

Integrating with respect to η and using the quasi-stationary property (3.26) and Fubini's Theorem (recall that T and the process (R, Υ) are independent), we obtain

$$\begin{aligned} \frac{1}{\bar{T} - \underline{T}} \int_{\underline{T}}^{\bar{T}} dt e^{\lambda c t} \eta[f] &= \frac{1}{\bar{T} - \underline{T}} \int_{\underline{T}}^{\bar{T}} dt \mathbf{E}_\eta^\dagger[f(R_t, \Upsilon_t); t < \mathbf{k}] \\ &= \mathbf{E}_\eta^\dagger[f(R_T, \Upsilon_T); T < \mathbf{k}] \\ &\leq C' \int_D dy \int_V dv f(y, v). \end{aligned} \tag{3.79}$$

Since f was chosen arbitrarily, this proves that η admits a uniformly bounded density (from above) with respect to the Lebesgue measure on $D \times V$.

Finally, using the quasi-stationarity of η (3.26) and integrating inequality (3.44) with respect to η implies that (here the time t and the constants $k, C_\varepsilon, c_\varepsilon$ depend on ε as in inequality (3.45)), for all bounded measurable functions f on $D \times E$,

$$\begin{aligned} e^{\lambda c t} \int_{D \times V} f(x) \eta(dx) &= \mathbf{E}_\eta^\dagger[f(R_t, \Upsilon_t); t < \mathbf{k}] \\ &\geq \eta(D_\varepsilon \times V) C_\varepsilon c_\varepsilon^k \int_{D_\varepsilon \times V} f(z, w) dz dw. \end{aligned}$$

This implies that $\tilde{\varphi}$ is a.e. lower bounded by $e^{-\lambda c t} \eta(D_\varepsilon \times V) C_\varepsilon c_\varepsilon^k$ on $D_\varepsilon \times V$. Since this inequality can be proved for any $\varepsilon > 0$ small enough, one deduces that, on any subset $D_\varepsilon \times V$ with $\varepsilon > 0$ and hence on any compactly embedded subset of $D \times V$, $\tilde{\varphi}$ is a.e. uniformly bounded away from zero. \square

3.8 Proof of Theorem 9

There are three main steps to the proof. The first is to characterise the law of transitions of the Markov process (X, \mathbb{P}^φ) , defined in the change of measure (3.21); note that the latter ensures the Markov property is preserved. The second step is to show that they agree with those of $(X^\varphi, \tilde{\mathbb{P}}^\varphi)$. The third step is to show that $(X^\varphi, \tilde{\mathbb{P}}^\varphi)$ is Markovian. Together these three imply the statement of the theorem.

Step 1. Next we look at the multiplicative semigroup which characterises uniquely the transitions of $(X^\varphi, \mathbb{P}^\varphi)$ (cf. [23, 24, 25])

$$u_t^\varphi[g](r, v) := \mathbb{E}_{\delta_{(r,v)}} \left[W_t \prod_{i=1}^{N_t} g(R_t^i, \Upsilon_t^i) \right] = \mathbb{E}_{\delta_{(r,v)}} \left[e^{-\lambda_* t} \frac{\langle \varphi, X_t \rangle}{\varphi(r, v)} e^{\langle \log g, X_t \rangle} \right] \quad (3.80)$$

for $t \geq 0$ and $g \in L_\infty^+(D \times V)$ which is uniformly bounded by unity. Note, we keep to the convention that an empty product is understood as 1, however we also define the empty inner product as zero (corresponding to all functions scoring zero when particles arrive at the cemetery state $\{\dagger\}$). As such, if we are to extend the domain of test functions in the product to include the cemetery state $\{\dagger\}$, we need to insist on the default value $g(\{\dagger\}) = 1$; see [23, 24, 25].

We start in the usual way by splitting the expectation in the second equality of (3.80) according to whether a scattering or fission event occurs. (The reader may wish to recall the

role of the quantities $\sigma_{\mathbf{s}}$, $\sigma_{\mathbf{f}}$, $\sigma = \sigma_{\mathbf{s}} + \sigma_{\mathbf{f}}$, $\pi_{\mathbf{s}}$ and $\sigma_{\mathbf{f}}$ in (3.1). We get

$$\begin{aligned}
u_t^\varphi[g](r, v) &= g(r + v(t \wedge \kappa_{(r,v)}^D), v) \frac{\varphi(r + vt, v)}{\varphi(r, v)} e^{-\int_0^t \lambda_* + \sigma(r+vs, v) ds} \mathbf{1}_{\{t < \kappa_{(r,v)}^D\}} \\
&+ \int_0^{t \wedge \kappa_{(r,v)}^D} \sigma_{\mathbf{s}}(r + vs, v) e^{-\int_0^s \lambda_* + \sigma(r+vl, v) dl} \frac{\varphi(r + vs, v)}{\varphi(r, v)} \\
&\quad \int_V u_{t-s}^\varphi[g](r + vs, v') \frac{\varphi(r + vs, v')}{\varphi(r + vs, v)} \pi_{\mathbf{s}}(r + vs, v, v') dv' \\
&+ \int_0^{t \wedge \kappa_{(r,v)}^D} \sigma_{\mathbf{f}}(r + vs, v) e^{-\int_0^s \lambda_* + \sigma(r+vl, v) dl} \frac{\varphi(r + vs, v)}{\varphi(r, v)} \\
&\quad \mathcal{E}_{(r+vs, v)} \otimes \mathbb{E}_{\delta_{(r,v)}} \left[\sum_{i=1}^N \frac{\varphi(r + vs, v_i)}{\varphi(r + vs, v)} W_{t-s}^i(r + vs, v_i) \prod_{j=1}^N e^{\langle \log g, X_{t-s}^j(r+vs, v_i) \rangle} \right], \tag{3.81}
\end{aligned}$$

where, for $r \in D$, $v \in V$, $W^i(r, v)$ and $X^i(r, v)$ are independent copies of the pair W and X under $\mathbb{P}_{\delta_{(r,v)}}$. Note that the first term on the right-hand side of (3.81) contains includes $g(r + v(t \wedge \kappa_{(r,v)}^D), v)$ to account for the fact that $g(\{\dagger\}) = 1$. Before developing the right-hand side above any further, we need to make two additional observations and to introduce some more notation.

The first observation is that, since W is a martingale, by sampling at the time of the first scattering event, fission event or when it leaves the domain D , whichever happens first, thanks to Doob's Optional Sampling Theorem, its mean must remain equal to 1 and we get the functional equation

$$\begin{aligned}
&\varphi(r, v) \\
&= \varphi(r + vt, v) e^{-\int_0^t \lambda_* + \sigma(r+vs, v) ds} \mathbf{1}_{\{t < \kappa_{(r,v)}^D\}} \\
&+ \int_0^{t \wedge \kappa_{(r,v)}^D} e^{-\int_0^s \lambda_* + \sigma(r+vl, v) dl} \sigma_{\mathbf{f}}(r + vs, v) \int_V \frac{\varphi(r + vs, v')}{\varphi(r + vs, v)} \pi_{\mathbf{f}}(r + vs, v, v') \varphi(r + vs, v) ds \\
&+ \int_0^{t \wedge \kappa_{(r,v)}^D} e^{-\int_0^s \lambda_* + \sigma(r+vl, v) dl} \sigma_{\mathbf{s}}(r + vs, v) \int_V \frac{\varphi(r + vs, v')}{\varphi(r + vs, v)} \pi_{\mathbf{s}}(r + vs, v, v') \varphi(r + vs, v) ds \tag{3.82}
\end{aligned}$$

for $r \in D$, $v \in V$. Now appealing to Lemma 1.2, Chapter 4 in [16], to treat the last two terms of (3.82) as potentials, with a little bit of algebra we can otherwise write the above as

$$\varphi(r, v) = \varphi(r + vt, v) \exp \left\{ \int_0^t \frac{(\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}} - \lambda_* \mathbf{I}) \varphi(r + vs, v)}{\varphi(r + vs, v)} ds \right\},$$

for $t < \kappa_{r,v}^D$, where \mathbf{I} is the identity operator, which is to say, for $t < \kappa_{r,v}^D$,

$$\begin{aligned} & \frac{\varphi(r+vt, v)}{\varphi(r, v)} e^{-\int_0^t \lambda_* + \sigma(r+vs, v) ds} \\ &= \exp \left\{ - \int_0^t \frac{(\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}} + \sigma \mathbf{I}) \varphi(r+vs, v)}{\varphi(r+vs, v)} ds \right\}. \end{aligned} \quad (3.83)$$

Our second observation pertains to the manipulation of the expectation on the right-hand side of (3.81). Define for $g \in L_\infty^+(D \times V)$, $(r, v) \in D \times V$ and $t \geq 0$,

$$u_t[g](r, v) := \mathbb{E}_{\delta_{(r,v)}} \left[\prod_{i=1}^{N_t} g(R_t^i, \Upsilon_t^i) \right] \quad (3.84)$$

We have that for all $(r, v) \in D \times V$,

$$\begin{aligned} & \sigma_{\mathbf{f}}(r, v) \mathcal{E}_{(r+vs, v)} \otimes \mathbb{E}_{\delta_{(r,v)}} \left[\sum_{i=1}^N \frac{\varphi(r, v_i)}{\varphi(r, v)} W_{t-s}^i(r, v_i) \prod_{j=1}^N e^{\langle \log g, X_{t-s}^j(r, v_i) \rangle} \right] \\ &= \sigma_{\mathbf{f}}(r, v) \mathcal{E}_{(r+vs, v)} \otimes \mathbb{E}_{\delta_{(r,v)}} \left[\frac{\langle \varphi, \mathcal{Z} \rangle}{\varphi(r, v)} \sum_{i=1}^N \frac{\varphi(r, v_i)}{\langle \varphi, \mathcal{Z} \rangle} W_{t-s}^i(r, v_i) e^{\langle \log g, X_{t-s}^i(r, v_i) \rangle} \prod_{\substack{j=1 \\ i \neq j}}^N e^{\langle \log g, X_{t-s}^j(r, v_i) \rangle} \right] \\ &= \sigma_{\mathbf{f}}(r, v) \frac{\mathcal{E}_{(r,v)}[\langle \varphi, \mathcal{Z} \rangle]}{\varphi(r, v)} \mathcal{E}_{(r,v)} \left[\frac{\langle \varphi, \mathcal{Z} \rangle}{\mathcal{E}_{(r,v)}[\langle \varphi, \mathcal{Z} \rangle]} \sum_{i=1}^N \frac{\varphi(r, v_i)}{\langle \varphi, \mathcal{Z} \rangle} u_{t-s}^\varphi[g](r, v_i) \prod_{\substack{j=1 \\ i \neq j}}^N u_{t-s}[g](r, v_j) \right] \\ &= G_{\mathbf{f}}^\varphi[u_{t-s}^\varphi[g], u_{t-s}[g]](r, v) + \frac{(\overleftarrow{\mathbf{F}} + \sigma_{\mathbf{f}} \mathbf{I}) \varphi(r, v)}{\varphi(r, v)} u_{t-s}^\varphi[g](r, v), \end{aligned} \quad (3.85)$$

where in the penultimate equality we have taken expectations conditional on the fission event and

$$\begin{aligned} G_{\mathbf{f}}^\varphi[f, g](r, v) &:= \frac{(\overleftarrow{\mathbf{F}} + \sigma_{\mathbf{f}} \mathbf{I}) \varphi(r, v)}{\varphi(r, v)} \mathcal{E}_{(r,v)}^\varphi \left[\sum_{i=1}^N \frac{\varphi(r, v_i)}{\langle \varphi, \mathcal{Z} \rangle} f(r, v_i) \prod_{\substack{j=1 \\ i \neq j}}^N g(r, v_j) \right] \\ &\quad - \frac{(\overleftarrow{\mathbf{F}} + \sigma_{\mathbf{f}} \mathbf{I}) \varphi(r, v)}{\varphi(r, v)} f(r, v) \end{aligned} \quad (3.86)$$

for $f, g \in L_\infty^+(D \times V)$, which are uniformly bounded by unity, and for $r \in D$, $v \in V$, where we recall that $\mathcal{P}_{(r,v)}^\varphi$ was defined in (3.23). Note in particular that

$$\frac{\mathcal{E}_{(r,v)}[\langle \varphi, \mathcal{Z} \rangle]}{\varphi(r, v)} = \int_V \frac{\varphi(r, v')}{\varphi(r, v)} \pi_{\mathbf{f}}(r, v, v') dv' = \frac{(\overleftarrow{\mathbf{F}} + \sigma_{\mathbf{f}} \mathbf{I}) \varphi(r, v)}{\sigma_{\mathbf{f}}(r, v) \varphi(r, v)}. \quad (3.87)$$

We will also make use of the notation

$$G_{\mathbf{f}}[f](r, v) = \sigma_{\mathbf{f}}(r, v) \mathcal{E}_{(r, v)} \left[\prod_{j=1}^N g(r, v_j) - g(r, v) \right], \quad (3.88)$$

for $r \in D$, $v \in V$ and $g \in L_{\infty}^+(D \times V)$, which is uniformly bounded by unity. Recall that the empty product in the definition (3.84) is defined as unity.

In a similar manner to (3.81) we can break the expectation over the event of scattering or fission in (3.84), which defines $u_t[g]$, to see that the operator $G_{\mathbf{f}}$ appears in the decomposition

$$u_t[g](r, v) = \hat{\mathbf{U}}_t[g] + \int_0^t \mathbf{U}_s[\overleftarrow{\mathbf{S}} u_{t-s}[g] + G_{\mathbf{f}}[u_{t-s}[g]]] ds, \quad t \geq 0, \quad (3.89)$$

for $g \in L_{\infty}^+(D \times V)$, which is uniformly bounded by unity. Here, we have adjusted the definition of the semigroup \mathbf{U} to

$$\hat{\mathbf{U}}_t[g](r, v) = g(r + v(t \wedge \kappa_{r, v}^D), v), t \geq 0, r \in D, v \in V. \quad (3.90)$$

Now returning to (3.81) with the above observations and definitions in hand, whilst again appealing to Lemma 1.2, Chapter 4 in [16], we have

$$\begin{aligned} u_t^{\varphi}[g](r, v) &= g(r + v(t \wedge \kappa_{r, v}^D), v) \\ &+ \int_0^{t \wedge \kappa_{r, v}^D} \sigma_{\mathbf{s}}(r + vs, v) \int_V u_{t-s}^{\varphi}[g](r + vs, v') \frac{\varphi(r + vs, v')}{\varphi(r + vs, v)} \pi_{\mathbf{s}}(r + vs, v, v') dv' ds \\ &+ \int_0^{t \wedge \kappa_{r, v}^D} G_{\mathbf{f}}^{\varphi}[u_{t-s}^{\varphi}[g], u_{t-s}[g]](r + vs, v) + \frac{(\overleftarrow{\mathbf{F}} + \sigma_{\mathbf{f}} \mathbf{I}) \varphi(r + vs, v)}{\varphi(r + vs, v)} u_{t-s}^{\varphi}[g](r + vs, v) ds \\ &- \int_0^{t \wedge \kappa_{r, v}^D} \frac{(\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}} + \sigma \mathbf{I}) \varphi(r + vs, v)}{\varphi(r + vs, v)} u_{t-s}^{\varphi}[g](r + vs, v) ds \end{aligned} \quad (3.91)$$

$$= \hat{\mathbf{U}}_t[g] + \int_0^t \mathbf{U}_s \left[\overleftarrow{\mathbf{S}}_{\varphi} u_{t-s}^{\varphi}[g] \right] ds + \int_0^t \mathbf{U}_s \left[G_{\mathbf{f}}^{\varphi}[u_{t-s}^{\varphi}[g], u_{t-s}[g]] \right] ds, \quad (3.92)$$

where where

$$\overleftarrow{\mathbf{S}}_{\varphi} f(r, v) := \int_V [f(r, v') - f(r, v)] \sigma_{\mathbf{s}}(r, v) \frac{\varphi(r, v')}{\varphi(r, v)} \pi_{\mathbf{s}}(r, v, v') dv'$$

on $D \times V$ and otherwise equal to zero,

Step 2. Define

$$\tilde{u}_t^{\varphi}[g](r, v) = \tilde{\mathbb{E}}_{\delta_{(r, v)}}^{\varphi} \left[\prod_{i=1}^{N_t} g(R_t^i, \Upsilon_t^i) \right], \quad t \geq 0, \quad (3.93)$$

for $g \in L_{\infty}^+(D \times V)$, where $\{(R_t^i, \Upsilon_t^i) : i = 1, \dots, N_t\}$ are the physical configurations of the particles alive in the system at time $t \geq 0$ in X^{φ} .

By conditioning \tilde{u}_t^φ on the first time a scattering of fission event occurs, it is a straightforward exercise to show that it also solves (3.92). For the sake of brevity, we leave this as an exercise to the reader as the arguments are similar to those mentioned previously. In order to show that (3.92) has a unique solution, we consider $v_t^\varphi[g] := \varphi u_t^\varphi[g]$ and $\tilde{v}_t^\varphi[g] := \varphi \tilde{u}_t^\varphi[g]$. Since u_t^φ and \tilde{u}_t^φ both satisfy (3.91), applying [16, Lemma 1.2, Chapter 4] along with (3.83), it is straightforward to show that v_t^φ and \tilde{v}_t^φ both satisfy

$$\begin{aligned} v_t[g](r, v) &= g(r + v(t \wedge \kappa_{(r,v)}^D), v) \varphi(r + vt, v) e^{-\int_0^t \lambda_* + \sigma(r+vs, v) ds} \mathbf{1}_{(t < \kappa_{(r,v)}^D)} \\ &\quad + \int_0^{t \wedge \kappa_{(r,v)}^D} e^{-\int_0^s \lambda_* + \sigma(r+vl, v) dl} \mathbf{U}_s[(\tilde{\mathbf{S}} + \sigma_{\mathbf{s}}) v_{t-s}[g]] ds \\ &\quad + \int_0^{t \wedge \kappa_{(r,v)}^D} e^{-\int_0^s \lambda_* + \sigma(r+vl, v) dl} \mathbf{U}_s[\tilde{G}_{\mathbf{f}}[v_{t-s}[g], u_{t-s}[g]] + \sigma_{\mathbf{f}} v_{t-s}[g]] ds, \end{aligned} \quad (3.94)$$

where

$$\tilde{G}_{\mathbf{f}}[f, g](r, v) = \sigma_{\mathbf{f}}(r, v) \left\{ \mathcal{E}_{(r,v)} \left[\sum_{i=1}^N f(r, v_i) \prod_{\substack{j=1 \\ j \neq i}}^N g(r, v_j) \right] - f(r, v) \right\}.$$

Due to the assumptions (H1) and (H4), an application of Grönwall's inequality implies uniqueness of (3.94), which in turn implies uniqueness of (3.92). We leave this as an exercise to the reader as it is a relatively standard computation and very similar to the calculations given in [10].

Step 3. We start by noting that the joint process $(X^\varphi, (R^\varphi, \Upsilon^\varphi))$ is, by construction, Markovian under $\tilde{\mathbb{P}}^\varphi$, we thus need to show that the marginalisation of the coupled system to just X^φ retains the Markov property. We do this by showing that for $f \in L_\infty^+(D \times V)$, $\mu \in \mathcal{M}(D \times V)$ and $(r, v) \in D \times V$,

$$\tilde{\mathbb{E}}_\mu^\varphi [f(R_t^\varphi, \Upsilon_t^\varphi) | X_t^\varphi] = \frac{\langle f \varphi, X_t^\varphi \rangle}{\langle \varphi, X_t^\varphi \rangle}, \quad t \geq 0. \quad (3.95)$$

This says that knowing X_t^φ only allows one to construct the law of $(R_t^\varphi, \Upsilon_t^\varphi)$ through an empirical distribution using φ . Hence, for $g \in L_\infty^+(D \times V)$ which is bounded by unity and $\mu \in \mathcal{M}(D \times V)$,

$$\begin{aligned} \tilde{\mathbb{E}}_\mu^\varphi \left[e^{\langle \log g, X_{t+s}^\varphi \rangle} | \mathcal{F}_t \right] &= \tilde{\mathbb{E}}_\mu^\varphi \left[\sum_{i=1}^{N_t} \frac{\varphi(R_t^i, \Upsilon_t^i)}{\langle \varphi, X_t^\varphi \rangle} \tilde{\mathbb{E}}_{\mu', (r,v)}^\varphi \left[e^{\langle \log g, X_s^\varphi \rangle} \right]_{\mu' = X_t^\varphi, (r,v)(R_t^i, \Upsilon_t^i)} \right] \\ &= \tilde{\mathbb{E}}_\mu^\varphi \left[\tilde{\mathbb{E}}_{\mu'}^\varphi \left[e^{\langle \log g, X_s^\varphi \rangle} \right]_{\mu' = X_t^\varphi} \right], \end{aligned}$$

where we have written $X_t^\varphi = \sum_{i=1}^{N_t} \delta_{(R_t^i, \Upsilon_t^i)}$, and thus the Markov property of $X^\varphi, \tilde{\mathbb{P}}^\varphi$ follows.

We are thus left with proving (3.95) to complete this step. To do so we note that it suffices to show that for $f, g \in L_\infty^+(D \times V)$ bounded by unity, $\mu \in \mathcal{M}(D \times V)$ and $(r, v) \in D \times V$,

$$\tilde{\mathbb{E}}_\mu^\varphi \left[f(R_t^\varphi, \Upsilon_t^\varphi) e^{\langle \log g, X_t^\varphi \rangle} \right] = \tilde{\mathbb{E}}_\mu^\varphi \left[\frac{\langle f\varphi, X_t^\varphi \rangle}{\langle \varphi, X_t^\varphi \rangle} e^{\langle \log g, X_t^\varphi \rangle} \right], \quad t \geq 0. \quad (3.96)$$

On the left-hand side of (3.96), we have

$$\begin{aligned} & \tilde{\mathbb{E}}_\mu^\varphi \left[f(R_t^\varphi, \Upsilon_t^\varphi) e^{\langle \log g, X_t^\varphi \rangle} \right] \\ &= \tilde{\mathbb{E}}_\mu^\varphi \left[\tilde{\mathbb{E}}_\mu^\varphi \left[f(R_t^\varphi, \Upsilon_t^\varphi) e^{\langle \log g, X_t^\varphi \rangle} \mid R_t^\varphi, \Upsilon_t^\varphi \right] \right] \\ &= \sum_{k=1}^n \frac{\varphi(r_k, v_k)}{\langle \varphi, \mu \rangle} \tilde{\mathbb{E}}_{\delta_{(r_k, v_k)}}^\varphi \left[f(R_t^\varphi, \Upsilon_t^\varphi) \prod_{i \geq 1: T_i \leq t} \prod_{j=1}^{N^i} \frac{u_{t-T_i}[g](R_{T_i}^\varphi, v_j^i)}{u_{t-T_i}[g](R_{T_i}^\varphi, \Upsilon_{T_i}^\varphi)} \right] \end{aligned}$$

where $\mu = \sum_{k=1}^n \delta_{(r_k, v_k)}$, T_i , $i \geq 1$ are the times of fission along the spine at which point N^i particles are issued at v_j^i , $j = 1 \cdots, N^i$. On the right-hand side of (3.96), we may appeal to Step 1 and Step 2 to deduce that

$$\begin{aligned} \tilde{\mathbb{E}}_\mu^\varphi \left[\frac{\langle f\varphi, X_t^\varphi \rangle}{\langle \varphi, X_t^\varphi \rangle} e^{\langle \log g, X_t^\varphi \rangle} \right] &= e^{-\lambda_* t} \mathbb{E}_\mu \left[\frac{\langle f\varphi, X_t \rangle}{\langle \varphi, \mu \rangle} e^{\langle \log g, X_t \rangle} \right] \\ &= \sum_{i=1}^n \frac{\varphi(r_i, v_i)}{\langle \varphi, \mu \rangle} e^{-\lambda_* t} \mathbb{E}_{\delta_{(r_i, v_i)}} \left[\frac{\langle f\varphi, X_t \rangle}{\varphi(r_i, v_i)} e^{\langle \log g, X_t \rangle} \right]. \end{aligned}$$

The proof of this final step is thus complete as soon as we can show that

$$\tilde{\mathbb{E}}_{\delta_{(r_k, v_k)}}^\varphi \left[f(R_t^\varphi, \Upsilon_t^\varphi) \prod_{1 \leq i: T_i \leq t} \prod_{j=1}^{N^i} \frac{u_{t-T_i}[g](R_{T_i}^\varphi, v_j^i)}{u_{t-T_i}[g](R_{T_i}^\varphi, \Upsilon_{T_i}^\varphi)} \right] = e^{-\lambda_* t} \mathbb{E}_{\delta_{(r_i, v_i)}} \left[\frac{\langle f\varphi, X_t \rangle}{\varphi(r_i, v_i)} e^{\langle \log g, X_t \rangle} \right] \quad (3.97)$$

To this end, we note that splitting the expectation on the right-hand of (3.97) side at either a scattering or fission event results in a calculation that is almost identical to the one above that concludes with (3.92). More precisely, the expectation on the right-hand side solves (3.92) albeit the role of $\hat{U}_t[g]$ is replaced by $U_t[f g]$. Similarly splitting the expectation on the left-hand side of (3.97) also results in a solution to (3.92) (with the aforementioned adjustment). The uniqueness of (3.92), with \hat{U} replaced by U , follows from the same arguments and hence the equality in (3.97) now follows, as required. \square

3.9 Proof of Lemma 5

The fact that the spine is Markovian is immediate from its definition of $(R^\varphi, \Upsilon^\varphi)$. Indeed, once given its initial configuration, it evolves as the NRW associated to the rate $\varphi^{-1}(r, v) \sigma_{\mathfrak{s}}(r, v) \varphi(r, v') \pi_{\mathfrak{s}}(r, v, v')$. Moreover, when in configuration $(r, v) \in D \times V$, at rate $\varphi(r, v)^{-1} (\bar{\mathbf{F}} + \sigma_{\mathfrak{f}} \mathbf{I}) \varphi(r, v)$, it experiences an additionally scattering with new velocity v' , with distribution

$$\mathcal{E}_{(r, v)} \left[\frac{\langle \varphi, \mathcal{Z} \rangle}{\mathcal{E}_{(r, v)}[\langle \varphi, \mathcal{Z} \rangle]} \frac{\langle \varphi \mathbf{1}_{(\cdot \in dv')} \rangle}{\langle \varphi, \mathcal{Z} \rangle} \right] = \frac{\sigma_{\mathfrak{f}}(r, v)}{(\bar{\mathbf{F}} + \sigma_{\mathfrak{f}} \mathbf{I}) \varphi(r, v)} \varphi(r, v') \pi_{\mathfrak{f}}(r, v, v') dv',$$

for $v' \in V$, where we have used (3.87). The total scatter rate is thus

$$\begin{aligned}
& \sigma_{\mathbf{s}}(r, v) \frac{\varphi(r, v')}{\varphi(r, v)} \pi_{\mathbf{s}}(r, v, v') + \sigma_{\mathbf{f}}(r, v) \frac{\varphi(r, v')}{\varphi(r, v)} \pi_{\mathbf{f}}(r, v, v') dv' \\
&= \alpha(r, v) \frac{\varphi(r, v')}{\varphi(r, v)} \pi(r, v, v') \\
&= \alpha^\varphi(r, v) \pi^\varphi(r, v, v')
\end{aligned} \tag{3.98}$$

as required.

For the second statement, write

$$\psi_t^\varphi[g](r, v) := \mathbf{E}_{(r, v)} \left[e^{-\lambda_* t + \int_0^t \beta(R_s, \Upsilon_s) ds} \frac{\varphi(R_t, \Upsilon_t)}{\varphi(r, v)} g(R_t, \Upsilon_t) \mathbf{1}_{(t < \tau^D)} \right]. \tag{3.99}$$

By conditioning the expectation on the right-hand side on the first scattering event we have, for $t \geq 0$, $r \in D$ and $v \in V$,

$$\begin{aligned}
\psi_t^\varphi[g](r, v) &= e^{-\lambda_* t + \int_0^t \beta(r + \nu \ell, v) - \alpha(r + \nu \ell, v) d\ell} \frac{\varphi(r + \nu t, v)}{\varphi(r, v)} g(r + \nu t, v) \mathbf{1}_{(t < \kappa_{r, v}^D)} \\
&+ \int_0^{t \wedge \kappa_{r, v}^D} \alpha(r + \nu s, v) e^{-\lambda_* s + \int_0^s \beta(r + \nu \ell, v) - \alpha(r + \nu \ell, v) d\ell} \frac{\varphi(r + \nu s, v)}{\varphi(r, v)} \\
&\quad \int_V \psi_{t-s}^\varphi[g](r + \nu s, v') \frac{\varphi(r + \nu s, v')}{\varphi(r + \nu s, v)} \pi(r + \nu s, v, v') dv' ds
\end{aligned} \tag{3.100}$$

Now appealing to (3.83), then using the standard trick of replacing the role of an additive potential by the role of a multiplicative potential in such semigroup evolutions, see e.g. Lemma 1.2, Chapter 4 in [16], and noting (3.99) we get

$$\psi_t^\varphi[g](r, v) = \mathbf{U}_t[g](r, v) + \int_0^t \mathbf{U}_s[(\tilde{\mathbf{L}}_\varphi + \alpha^\varphi \mathbf{I}) \psi_{t-s}^\varphi[g] - (\varphi^{-1}(\tilde{\mathbf{S}} + \tilde{\mathbf{F}})\varphi + (\beta - \alpha)\mathbf{I}) \psi_{t-s}^\varphi[g]](r, v) ds \tag{3.101}$$

where

$$\tilde{\mathbf{L}}_\varphi f(r, v) = \alpha^\varphi(r, v) \int_V [f(r, v') - f(r, v)] \pi^\varphi(r, v, v') dv', \tag{3.102}$$

for $f \in L_\infty^+(D \times V)$. Referring to (3.7), (3.8) and (3.15), we note that

$$\begin{aligned}
& \alpha^\varphi(r, v) - \frac{(\tilde{\mathbf{S}} + \tilde{\mathbf{F}})\varphi(r, v)}{\varphi(r, v)} - \beta(r, v) + \alpha(r, v) \\
&= \int_V \alpha(r, v) \frac{\varphi(r, v')}{\varphi(r, v)} \pi(r, v, v') dv' \\
&\quad - \int_V \frac{\varphi(r, v')}{\varphi(r, v)} (\sigma_{\mathbf{s}}(r, v) \pi_{\mathbf{s}}(r, v, v') + \sigma_{\mathbf{f}}(r, v) \pi_{\mathbf{f}}(r, v, v')) dv' - \sigma_{\mathbf{s}}(r, v) - \sigma_{\mathbf{f}}(r, v) \\
&\quad - \sigma_{\mathbf{f}}(r, v) \left(\int_V \pi_{\mathbf{f}}(r, v, v') dv' - 1 \right) + \sigma_{\mathbf{s}}(r, v) + \sigma_{\mathbf{f}}(r, v) \int_V \pi_{\mathbf{f}}(r, v, v') dv' \\
&= 0.
\end{aligned}$$

Hence (3.101) reduces to the somewhat simpler recurrence equation

$$\psi_t^\varphi[g](r, v) = \mathbf{U}_t[g](r, v) + \int_0^t \mathbf{U}_s[\tilde{\mathbf{L}}_\varphi \psi_{t-s}^\varphi[g]](r, v) ds, \quad t \geq 0,$$

where we recall that $\tilde{\mathbf{L}}_\varphi$ was defined in (3.102). This is nothing more than the mild equation for the semigroup evolution $\mathbf{E}_{(r,v)}^\varphi[g(R_t, \Upsilon_t)]$, $t \geq 0$, which has a unique bounded solution from the usual Grönwall arguments. Note that when $g = 1$, we see the solution is 1. This, together with the Markov property implies that the right-hand side of (3.25) is a martingale. Moreover, it follows that the martingale change of measure in (3.25) describes law of the $\alpha^\varphi \pi^\varphi$ -NRW.

The fact that $\psi_t^\varphi[1](r, v) = 1$ for all $r \in D, v \in V$, implies that $((R, \Upsilon), \mathbf{P}^\varphi)$ is conservative. Moreover,

$$\mathbf{P}_{(r,v)}^\varphi[g(R_t, \Upsilon_t)] = \psi_t^\varphi[g](r, v) = e^{-\lambda_* t} \frac{\psi_t[g\varphi](r, v)}{\varphi(r, v)}, \quad r \in D, v \in V.$$

where $g \in L_\infty^+(D \times V)$; cf. (3.99). Hence $\lim_{t \rightarrow \infty} \mathbf{P}_{(r,v)}^\varphi[g(R_t, \Upsilon_t)] = \langle g, \tilde{\varphi}\varphi \rangle$ for all $g \in L_\infty^+(D \times V)$. In other words, $\tilde{\varphi}\varphi$ is the density of the stationary distribution of (R, Υ) under \mathbb{P}^φ . \square

3.10 Proof of Theorem 8

The proof we offer here is a variant of a standard one, which has been used to analyse the convergence of many analogous martingales in the setting of different spatial branching processes. We mention [32], [40], [3] and [17] to name but a few of the contexts with similar results.

In the case that $\lambda_* < 0$ and $\lambda_* > 0$, we need (H3) to ensure that the NBP can undergo fission. In the setting $\lambda_* = 0$ we need the stricter condition (H3)* for technical reasons in the proof to ensure a minimal rate of reproduction.

A standard measure theoretic result (cf. p. 242 of [15]) tells us that the martingale change of measure in (3.21) is uniformly integrable if and only if

$$\mathbb{P}_{\delta_{(r,v)}}^\varphi \left(\limsup_{t \rightarrow \infty} W_t < \infty \right) = 1.$$

In the case that $\limsup_{t \rightarrow \infty} W_t = \infty$, \mathbb{P}^φ almost surely, we have $\mathbb{P}_{\delta_{(r,v)}}(W_\infty = 0) = 1$.

(i) Let us first deal with the case that $\lambda_* > 0$. To this end, let us define the sigma algebra $\mathcal{S} = \sigma(T_i, \mathcal{Z}_i : i \geq 0)$, where T_i , $i \geq 1$, are the times at which the spine undergoes fission and \mathcal{Z}_i , $i \geq 1$, are point processes on V that describe the velocities of fission offspring (i.e. whose law is given by the family (3.3) under the change of measure (3.23)). For convenience we will write $T_0 = 0$.

Appealing to the pathwise spine decomposition in Theorem 9, letting T_j denote the birth

times along the spine and N_j the number of offspring (excluding the spine itself), we can write

$$\begin{aligned}
\mathbb{E}_{\delta(r,v)}^\varphi[W_t] &= \mathbb{E}_{\delta(r,v)}^\varphi \left[\frac{e^{-\lambda_* t} \varphi(R_t, \Upsilon_t)}{\varphi(r, v)} \right] \\
&\quad + \mathbb{E}_{\delta(r,v)}^\varphi \left[\mathbb{E}_{\delta(r,v)}^\varphi \left[\sum_{j=1}^{\infty} e^{-\lambda_* T_j} \mathbf{1}_{(T_j \leq t)} \sum_{i=1}^{N_j} \varphi(R_{T_j}, v_i) W_{t-T_j}^j(R_{T_j}, v_i) \middle| \mathcal{S} \right] \right] \\
&= \mathbb{E}_{\delta(r,v)}^\varphi \left[\frac{e^{-\lambda_* t} \varphi(R_t, \Upsilon_t)}{\varphi(r, v)} + \sum_{j=1}^{\infty} e^{-\lambda_* T_j} \mathbf{1}_{(T_j \leq t)} \mathcal{E}_{(R_{T_j}, \Upsilon_{T_{j-1}})}^\varphi[\langle \varphi, \mathcal{Z}_j \rangle] \right] \\
&= \mathbb{E}_{\delta(r,v)}^\varphi \left[\frac{e^{-\lambda_* t} \varphi(R_t, \Upsilon_t)}{\varphi(r, v)} + \sum_{j=1}^{\infty} e^{-\lambda_* T_j} \mathbf{1}_{(T_j \leq t)} \frac{\mathcal{E}_{(R_{T_j}, \Upsilon_{T_{j-1}})}[\langle \varphi, \mathcal{Z} \rangle^2]}{\mathcal{E}_{(R_{T_j}, \Upsilon_{T_{j-1}})}[\langle \varphi, \mathcal{Z} \rangle]} \right] \tag{3.103}
\end{aligned}$$

where, for a given $(r, v) \in D \times V$, the process $W_s^j(r, v)$ is an independent copy of $(W_s, s \geq 0)$, under $\mathbb{P}_{\delta(r,v)}$ (and consequently has unit mean, which is also used above). Our objective is to prove that the right-hand side of (3.103) is finite. In that case, it will follow with the help of Fatou's Lemma that

$$\infty > \limsup_{t \rightarrow \infty} \mathbb{E}_{\delta(r,v)}^\varphi[W_t] \geq \mathbb{E}_{\delta(r,v)}^\varphi \left[\liminf_{t \rightarrow \infty} W_t \right]. \tag{3.104}$$

Recalling that W is a non-negative \mathbb{P} -martingale, it holds that $1/W$ is a non-negative \mathbb{P}^φ -supermartingale and thus its limit exists [21]. The conditional expectation in (3.104) ensures that $\liminf_{t \rightarrow \infty} W_t$ (and hence from the immediately preceding remarks $\limsup_{t \rightarrow \infty} W_t$) is $\mathbb{P}_{\delta(r,v)}^\varphi$ -almost surely finite.

To this end, we again recall the description of the pathwise spine decomposition in Theorem 9 and note that fission along the spine occurs at the accelerated rate $\varphi^{-1}(\bar{\mathbf{F}} + \sigma_{\mathbf{f}} \mathbf{I})\varphi$. Hence (recalling the generic point process \mathcal{Z} defined in (3.3))

$$\begin{aligned}
&\mathbb{E}_{\delta(r,v)}^\varphi \left[\sum_{j=1}^{\infty} e^{-\lambda_* T_j} \mathbf{1}_{(T_j \leq t)} \frac{\mathcal{E}_{(R_{T_j}, \Upsilon_{T_{j-1}})}[\langle \varphi, \mathcal{Z} \rangle^2]}{\mathcal{E}_{(R_{T_j}, \Upsilon_{T_{j-1}})}[\langle \varphi, \mathcal{Z} \rangle]} \right] \\
&\leq (\|\varphi\|_\infty n_{\max})^2 \mathbb{E}_{\delta(r,v)}^\varphi \left[\sum_{j=1}^{\infty} e^{-\lambda_* T_j} \frac{1}{\mathcal{E}_{(R_{T_j}, \Upsilon_{T_{j-1}})}[\langle \varphi, \mathcal{Z} \rangle]} \right] \\
&= (\|\varphi\|_\infty n_{\max})^2 \mathbf{E}_{(r,v)}^\varphi \left[\int_0^\infty e^{-\lambda_* t} \frac{\sigma_{\mathbf{f}}(R_t, \Upsilon_t)}{\mathcal{E}_{(R_t, \Upsilon_t)}[\langle \varphi, \mathcal{Z} \rangle]} \int_V \frac{\varphi(R_t, v')}{\varphi(R_t, \Upsilon_t)} \pi_{\mathbf{f}}(R_t, \Upsilon_t, v') dv' dt \right] \\
&\leq \bar{\sigma}_{\mathbf{f}} (\|\varphi\|_\infty n_{\max})^2 \mathbf{E}_{(r,v)}^\varphi \left[\int_0^\infty \frac{e^{-\lambda_* t}}{\varphi(R_t, \Upsilon_t)} dt \right] \\
&\leq \bar{\sigma}_{\mathbf{f}} \frac{(\|\varphi\|_\infty n_{\max})^2}{\varphi(r, v)} \mathbf{E}_{(r,v)} \left[\int_0^\infty e^{-2\lambda_* t + \int_0^t \beta(R_s, \Upsilon_s) ds} \mathbf{1}_{(t < \tau^D)} dt \right] \\
&\leq \bar{\sigma}_{\mathbf{f}} (\|\varphi\|_\infty n_{\max})^2 \int_0^\infty e^{-2\lambda_* t} \frac{\psi_t[\mathbf{1}](r, v)}{\varphi(r, v)} dt, \tag{3.105}
\end{aligned}$$

where we have used (H4) in the first inequality, features of the spine decomposition for the first equality, (3.87) and (H1) in the second inequality, the change of measure (3.25) in the third

inequality and the semigroup (3.16) for the final line. Finally, note that since φ is uniformly bounded above, the contribution from the spine term in (3.103) is zero in the limit $t \rightarrow \infty$. Now using Theorem 7, it follows that

$$\limsup_{t \rightarrow \infty} \mathbb{E}_{\delta(r,v)}^\varphi [W_t] < \infty$$

as required.

(ii) Next, for the case $\lambda_* < 0$, it is easy to see that, on the event $\{T_j \leq t < T_{j+1}\}$,

$$W_t \geq e^{-\lambda_* t} \varphi(R_t, \Upsilon_{T_{j-1}})$$

which ensures that $\mathbb{P}_{\delta(r,v)}^\varphi(\limsup_{t \rightarrow \infty} W_t = \infty) = 1$ for all $(r, v) \in D \times V$ and hence $\mathbb{P}_{\delta(r,v)}(W_\infty = 0) = 1$.

(iii) Finally, for the case $\lambda_* = 0$, our aim is to show that, for each $r \in D, v \in V$,

$$\mathbb{P}_{\delta(r,v)}^\varphi(\limsup_{t \rightarrow \infty} W_t = \infty) = 1.$$

We do this by constructing a random sequence of times $(s_n : n \geq 0)$ such that $\limsup_{n \rightarrow \infty} W_{s_n} = \infty$ almost surely with respect to \mathbb{P}^φ .

Lemma 5 tells us that $\tilde{\varphi}$ is the density of the stationary distribution of (R, Υ) under \mathbf{P}^φ . Moreover, thanks to Theorem 7, the density $\tilde{\varphi}$ is a.e. uniformly bounded away from 0 on each $\Omega \subset\subset D \times V$. It follows that $\langle \mathbf{1}_\Omega, \varphi \tilde{\varphi} \rangle > 0$ for all $\Omega \subseteq D_\varepsilon \times V$ and that the spine R visits Ω infinitely often under \mathbf{P}^φ .

Fix $k \in \mathbb{N}$. We want to show that there is an $\Omega \subseteq D_\varepsilon \times V$ such that

$$\inf_{(r', v') \in \Omega} \mathbb{P}_{\delta(r', v')}^\varphi(X_\iota(D_\varepsilon \times V) \geq k) > 0, \quad (3.106)$$

where $\iota = 2\text{diam}(D)/\mathbf{v}_{\min}$ (note that ι is twice the time it would take a neutron to cross the equivalent of the diameter of D , when moving at minimal speed). To this end, fix $r \in D, v \in V$ and choose $\varepsilon > 0$ sufficiently small such that both $r \in D_\varepsilon := \{r \in D : \inf_{y \in \partial D} |r - y| > \varepsilon \mathbf{v}_{\max}\}$ and B (introduced in the assumption (H3)^{*}) is in D_ε . Then, define

$$\Omega = \{(r, v) \in D_\varepsilon \times V : \{r + vs : s \geq 0\} \cap B \neq \emptyset\}.$$

Write m for the smallest natural number such that $m(n_{\max} - 1) + 1 \geq k$. Recalling from Theorem 7 that $\inf_{r \in D, v, v' \in V} \alpha(r, v) \pi(r, v, v') > 0$, and taking account of the positivity properties of φ , we can lower bound the probability that, from any $(r, v) \in \Omega$, the spine can enter B . Moreover, on this event, due to (H3)^{*}, we can also lower bound the probability that the spine immigrates $n_{\max} - 1$ particles on m (evenly spaced in time) separate occasions, all of which are still inside of B by time ι . The strategy for doing so is to head into B from the given point of issue in Ω by travelling in a straight line within a small cone of possible velocities (which

would be guaranteed to happen within $\iota/2$ units of time), and then for neutrons to cycle around the perimeter of B in an annulus by scattering within a narrow cone of velocities each time; see Fig 3-2. As such we can provide the lower bound desired in (3.106). The technical details are tedious and left to the reader.

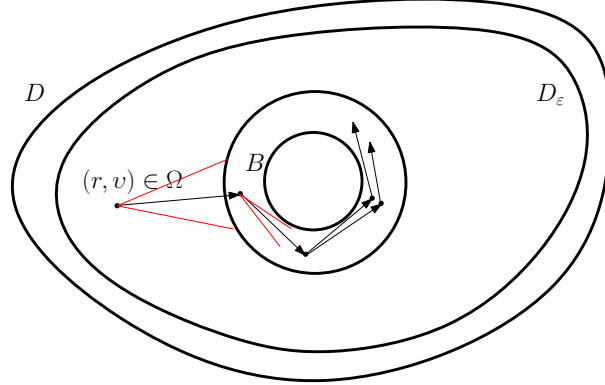


Figure 3-2: There is a uniform lower bound on the probability that the spine issued from $(r, v) \in \Omega$ heads directly into the annulus contained in B and subsequently immigrates $n_{\max} - 1$ neutrons on each of m separate occasions, which then cycle around the annulus, and all this is completed over the time horizon $\iota = 2\text{diam}(D)/v_{\min}$ elapses.

With (3.106) in hand, we can construct the sequence $(t_n : n \geq 0)$ by defining $t_0 = 0$ and subsequently, for $n \geq 1$,

$$t_n = \inf\{s > t_{n-1} + (10m \times \iota) : (R_s, \Upsilon_s) \in \Omega\}.$$

Note that since (R, Υ) visits Ω infinitely often under \mathbf{P}^φ we have that $t_n < \infty$, \mathbf{P}^φ -almost surely for $n \geq 0$, and $t_n \rightarrow \infty$, \mathbf{P}^φ -almost surely. By applying the strong Markov property at the sequence of times $(t_n, n \geq 0)$, it now follows from (3.106) that, in the spirit of a sequence of independent Bernoulli trials, $\limsup_{n \rightarrow \infty} X_{s_n}(D_\epsilon) \geq k$ almost surely with respect to \mathbb{P}^φ , where $s_n = t_n + (m \times \iota)$. Since the integer k can be chosen arbitrarily large, we also have that $\limsup_{n \rightarrow \infty} X_{s_n}(D_\epsilon) = \infty$ almost surely with respect to \mathbb{P}^φ .

As φ is uniformly bounded below away from 0 on $D_\epsilon \times V$ (see Theorem 7), it follows that

$$W_t \geq cX_t(D_\epsilon \times V), \quad t \geq 0,$$

for some constant $c > 0$. The analysis above, thus shows that $\mathbb{P}_{\delta(r,v)}^\varphi(\limsup_{t \rightarrow \infty} W_t = \infty) = 1$, as required, for each $r \in D$, $v \in V$.

□

3.11 Proof of Corollary 2

Doob's martingale inequality ensures that, for $\mu \in \mathcal{M}(D \times V)$

$$\mathbb{E}_\mu[(\sup_{t \geq 0} W_t)^2] \leq \liminf_{s \rightarrow \infty} 4\mathbb{E}_\mu[(W_s)^2].$$

Showing that the right-hand side above is finite is sufficient to obtain $L_2(\mathbb{P})$ convergence. Note, however, that $\mathbb{E}_\mu[(W_s)^2] = \mathbb{E}_\mu^\varphi[W_t]$, $t \geq 0$, and hence, from (3.104), the desired upper bound is proved. \square

Acknowledgements

This research was born out of a surprising connection that was made at the problem formulation “Integrative Think Tank” as part of the EPSRC Centre for Doctoral Training SAMBa in the summer of 2015. AEK and EH are indebted to Professor Paul Smith and Dr. Geoff Dobson from the ANSWERS modelling group at Wood for the extensive discussions at their offices in Dorchester as well as for giving permission to use these images in Figure 3-1, which were constructed with Wood nuclear software ANSWERS. We would also like to thank Alex Cox, Simon Harris and Minmin Wang for helpful discussions, as well as Jean Bertoin and Alex Watson for several interesting discussions on growth-fragmentation equations. Finally, we are extremely grateful to the assistance of an AE and an anonymous referee, which have helped us shape the presentation significantly.

Bibliography

- [1] G. Alsmeyer and A. Iksanov. A log-type moment result for perpetuities and its application to martingales in supercritical branching random walks. *Electron. J. Probab.*, 14:no. 10, 289–312, 2009.
- [2] G. I. Bell. On the stochastic theory of neutron transport. *Nuc. Sci. & Eng.*, 21:390–401, 1965.
- [3] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [4] J. Bertoin. On a Feynman-Kac approach to growth-fragmentation semigroups and their asymptotic behaviors. <https://arxiv.org/abs/1804.04905>, 2018.
- [5] J. Bertoin and A. R. Watson. A probabilistic approach to spectral analysis of growth-fragmentation equations. *J. Funct. Anal.*, 274(8):2163–2204, 2018.
- [6] J. D. Biggins. Martingale convergence in the branching random walk. *J. Appl. Probab.*, 14(1):25–37, 1977.
- [7] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q -process. *Probab. Theory Related Fields*, 164(1-2):243–283, 2016.
- [8] N. Champagnat and D. Villemonais. Uniform convergence to the q -process. *Electron. Commun. Probab.*, 22:7 pp., 2017.
- [9] A. M. G. Cox, S.C. Harris, A.E. Kyprianou, and M. Wang. Monte Carlo methods for the neutron transport equation. *Working document*, 2018.
- [10] A.M.G. Cox, S.C. Harris, E. Horton, and A.E. Kyprianou. Multi-species neutron transport equation. *J. Stat. Phys*, 176(2):425–455, 2019.
- [11] 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 à l'énergie atomique. Eyrolles, Paris, 1989.
- [12] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 6*. Springer-Verlag, Berlin, 1993. Evolution problems. II, With the collaboration of Claude Bardos, Michel Cessenat, Alain Kavenoky, Patrick Lascaux, Bertrand

Mercier, Olivier Pironneau, Bruno Scheurer and Rémi Sentis, Translated from the French by Alan Craig.

- [13] B. Davison and J. B. Sykes. *Neutron transport theory*. Oxford, at the Clarendon Press, 1957.
- [14] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4):1671–1691, 1995.
- [15] R. Durrett. *Probability: Theory and Examples*. Duxbury, Belmont, CA, 1996.
- [16] E. B. Dynkin. *Diffusions, superdiffusions and partial differential equations*, volume 50 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2002.
- [17] János Engländer and Andreas E. Kyprianou. Local extinction versus local exponential growth for spatial branching processes. *Ann. Probab.*, 32(1A):78–99, 2004.
- [18] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [19] R. Hardy and S. C. Harris. A spine approach to branching diffusions with applications to L^p -convergence of martingales. In *Séminaire de Probabilités XLII*, volume 1979 of *Lecture Notes in Math.*, pages 281–330. Springer, Berlin, 2009.
- [20] S.C. Harris, E. Horton, and A.E. Kyprianou. Stochastic analysis of the neutron transport equation ii: Almost sure growth. *Preprint*, 2019.
- [21] S.C. Harris and M.I. Roberts. Measure changes with extinction. *Stat Probabil Lett.*, 79(8):1129, 2009.
- [22] T. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)].
- [23] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. I. *J. Math. Kyoto Univ.*, 8:233–278, 1968.
- [24] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. II. *J. Math. Kyoto Univ.*, 8:365–410, 1968.
- [25] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. III. *J. Math. Kyoto Univ.*, 9:95–160, 1969.
- [26] K. Jörgens. An asymptotic expansion in the theory of neutron transport. *Comm. Pure Appl. Math.*, 11:219–242, 1958.

- [27] A. E. Kyprianou and A. Murillo-Salas. Super-Brownian motion: L^p -convergence of martingales through the pathwise spine decomposition. In *Advances in superprocesses and nonlinear PDEs*, volume 38 of *Springer Proc. Math. Stat.*, pages 113–121. Springer, New York, 2013.
- [28] B. Lapeyre, É. Pardoux, and R. Sentis. *Introduction to Monte-Carlo methods for transport and diffusion equations*, volume 6 of *Oxford Texts in Applied and Engineering Mathematics*. Oxford University Press, Oxford, 2003. Translated from the 1998 French original by Alan Craig and Fionn Craig.
- [29] J. Lehner. The spectrum of the neutron transport operator for the infinite slab. *J. Math. Mech.*, 11:173–181, 1962.
- [30] J. Lehner and G. M. Wing. On the spectrum of an unsymmetric operator arising in the transport theory of neutrons. *Comm. Pure Appl. Math.*, 8:217–234, 1955.
- [31] J. Lehner and G. M. Wing. Solution of the linearized Boltzmann transport equation for the slab geometry. *Duke Math. J.*, 23:125–142, 1956.
- [32] R. Lyons. A simple path to Biggins’ martingale convergence for branching random walk. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 245–249. Springer, New York, 1997.
- [33] S. Maire and D. Talay. On a Monte Carlo method for neutron transport criticality computations. *IMA J. Numer. Anal.*, 26(4):657–685, 2006.
- [34] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. in Appl. Probab.*, 24(3):542–574, 1992.
- [35] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.*, 25(3):487–517, 1993.
- [36] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.
- [37] M. Mokhtar-Kharroubi. *Mathematical topics in neutron transport theory*, volume 46 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co., Inc., River Edge, NJ, 1997. New aspects, With a chapter by M. Choulli and P. Stefanov.
- [38] I. Pázsit and L. Pál. *Neutron Fluctuations: A Treatise on the Physics of Branching Processes*. Elsevier, 2008.
- [39] A. Pazy and P. H. Rabinowitz. A nonlinear integral equation with applications to neutron transport theory. *Arch. Rational Mech. Anal.*, 32:226–246, 1969.
- [40] Z. Shi. *Branching random walks*, volume 2151 of *Lecture Notes in Mathematics, Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012, École d’Été de Probabilités de Saint-Flour*. Springer, 2015.

Glossary of some commonly used notation
(Th. = Theorem, a. = above, b. = below)

Notation	Description	Introduced
$(\psi_t, t \geq 0)$	Solution to mild NTE/NBP expectation semigroup	(3.5), (3.9)
D and V	Physical and velocity domain	§3.1
σ_s, σ_f and σ	Scatter, fission and total cross-sections	b. (3.1)
π_s and π_f	Scatter and fission kernels	b. (3.1)
\overleftarrow{S} and \overleftarrow{F}	Scatter and fission operators	(3.7), (3.8)
$\mathcal{P}_{(r,v)}$	Offspring law of X when parent at $(r, v) \in D \times V$	(3.4)
$((r_i, v_i), i = 1, \dots, N)$	Position and number of offspring of a family in X	a. (3.3)
(X, \mathbb{P}_μ)	NBP when issued from μ	§3.3
U_t	Linear advection semigroup	b. (3.6), (3.9)
G_f	Branching generator of (X, \mathbb{P}_μ)	(3.88)
$(u_t, t \geq 0)$	Non-linear semigroup of X	(3.84)
\hat{U}_t	Non-linear advection semigroup	(3.89), (3.90)
n_{\max}	Maximum number of neutrons in a fission event	b. (3.4), (H4)
λ_*, φ and $\tilde{\varphi}$	Leading eigenvalue, right- and left-eigenfunctions	Th. 7
$(W_t, t \geq 0)$	Additive martingale	(3.19)
<hr/>		
$((R, \Upsilon), \mathbf{P})$	Many-to-one NRW	Lemma 4
α and π	Scatter rate and kernel for many-to-one NRW	(3.11), (3.12)
τ^D	First exit time of $\alpha\pi$ -NRW from D	Lemma 4
$((R, \Upsilon), \mathbf{P}^\dagger)$	Killed $\alpha\pi$ -NRW	(3.17)
\mathbf{P}^\dagger	Semigroup of killed $\alpha\pi$ -NRW	(3.17)
β	Many-to-one potential	(3.15), (3.16)
\mathbf{k}	Killing time of $\alpha\pi$ -NRW	(3.18)
$(J_k, k \geq 1)$	Ordered jump times of killed $\alpha\pi$ -NRW	b. (3.30)
<hr/>		
$(X, \mathbb{P}_\mu^\varphi)$	NBP after change of measure with W	(3.21)
G_f^φ	Branching generator of $(X, \mathbb{P}_\mu^\varphi)$	(3.86), (3.92)
$(u_t^\varphi, t \geq 0)$	Non-linear semigroup of $(X, \mathbb{P}_\mu^\varphi)$	(3.80)
$(\psi_t^\varphi, t \geq 0)$	Linear semigroup of $(X, \mathbb{P}_\mu^\varphi)$	(3.99)
$(X^\varphi, \tilde{\mathbb{P}}_\mu^\varphi)$	Dressed spine when issued from configuration μ	(3.22)
$(\tilde{u}_t^\varphi, t \geq 0)$	Non-linear semigroup of $(X^\varphi, \tilde{\mathbb{P}}_\mu^\varphi)$	(3.93)
$((R^\varphi, \Upsilon^\varphi), \tilde{\mathbf{P}}^\varphi)$	Marginal of $\tilde{\mathbb{P}}^\varphi$ giving law of spine NRW	Th. 9, (3.25)
α^φ and π^φ	Scatter rate and kernel of auxiliary NRW	(3.24)

$\mathcal{P}_{(r,v)}^\varphi$	Scattering of velocities along the spine	(3.23)
\mathbf{P}^φ	Law of $\alpha^\varphi\pi^\varphi$ -NRW that agrees with $\tilde{\mathbf{P}}^\varphi$	a. (3.24)

Concluding remarks

In this chapter we considered the so-called neutron branching process and its linear semigroup, in order to analyse the neutron transport equation from a stochastic point of view. We also obtained a second stochastic representation, called the neutron random walk, via a many-to-one formula. Using theory from quasi-stationary distributions, we were able to analyse this latter process in order to obtain existence of the leading eigenvalue and associated left and right eigenfunctions. We then obtained a Perron-Fröbenius decomposition, which explicitly characterises the leading order behaviour of the branching system in terms of these leading eigenelements.

In the second half of this chapter, we looked at the different regimes of the branching process in accordance with the sign of the leading eigenvalue. We also obtained a spine decomposition, which we will exploit in the next chapter to obtain a strong law of large numbers result.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Stochastic Methods for the Neutron Transport Equation II: Almost sure growth			
Publication status (tick one)			
Draft manuscript	<input type="checkbox"/>	Submitted	<input type="checkbox"/>
		In review	<input checked="" type="checkbox"/>
		Accepted	<input type="checkbox"/>
		Published	<input type="checkbox"/>
Publication details (reference)	Under major review for Annals of Applied Probability: AAP1901-002 arxiv:1901.00220		
Copyright status (tick the appropriate statement)			
I hold the copyright for this material		<input type="checkbox"/>	Copyright is retained by the publisher, but I have been given permission to replicate the material here
			<input checked="" type="checkbox"/>
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	<p>*All research has been conducted in equal partnership with collaborators and supervisors. It is unwise to try and measure percentages of intellectual contribution as research scholarship lies as much in the escalation of ideas through mathematical discourse as it does with the seed of ideas themselves. That said, the mathematical content of this thesis is, as an entire piece of work, inextricably associated to its author through intellectual ownership.</p> <p>Formulation of ideas: The candidate played an integral and fully collaborative role in the formulation of ideas.</p> <p>Design of methodology: The candidate played an integral and fully collaborative role in the design of methodology.</p> <p>Experimental work: N/A</p> <p>Presentation of data in journal format: N/A</p> <p>*The wording in this box follows the advice and approval of my supervisor, Professor Kyprianou.</p>		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Emma Horton	Date	19/11/2019

Chapter 4

Skeleton decompositions and a strong law of large numbers

Simon C. Harris¹, Emma Horton², Andreas E. Kyprianou³.

Remark 7. *This article has been submitted using the name *Stochastic Methods for the Neutron Transport Equation II: Almost sure growth**

Abstract

The neutron transport equation (NTE) describes the flux of neutrons across a planar cross-section in an inhomogeneous fissile medium when the process of nuclear fission is active. Classical work on the NTE emerges from the applied mathematics literature in the 1950s through the work of R. Dautray and collaborators, [7, 8, 24]. The NTE also has a probabilistic representation through the semigroup of the underlying *physical process* when envisaged as a stochastic process; cf. [7, 22, 23, 25]. More recently, [6] and [17] have continued the probabilistic analysis of the NTE, introducing more recent ideas from the theory of spatial branching processes and quasi-stationary distributions. In this paper, we continue in the same vein and look at a fundamental description of stochastic growth in the supercritical regime. Our main result provides a significant improvement on the last known contribution to growth properties of the physical process in [25], bringing neutron transport theory in line with modern branching process theory such as [14, 12]. An important aspect of the proofs focuses on the use of a skeletal path decomposition, which we derive for general branching particle systems in the new context of non-local branching mechanisms.

¹Department of Statistics, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. Email: simon.harris@auckland.ac.nz

²Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: elh48@bath.ac.uk

³Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: a.kyprianou@bath.ac.uk

4.1 Introduction

In this article we continue our previous work in [17] and look in more detail at the stochastic analysis of the Markov process that lies behind the Neutron Transport Equation (NTE). We recall that the latter describes the flux, Ψ_t , at time $t \geq 0$, of neutrons across a planar cross-section in an inhomogeneous fissile medium (measured in number of neutrons per cm^2 per second). Neutron flux is described in terms of the configuration variables $(r, v) \in D \times V$, where $D \subseteq \mathbb{R}^3$ is (in general) a non-empty, smooth, open, bounded and convex domain such that ∂D has zero Lebesgue measure, and V is the velocity space, which is given by $V = \{v \in \mathbb{R}^3 : v_{\min} \leq |v| \leq v_{\max}\}$, where $0 < v_{\min} < v_{\max} < \infty$.

In its backwards form, the NTE is introduced as an integro-differential equation of the form

$$\begin{aligned} \frac{\partial}{\partial t} \psi_t(r, v) &= v \cdot \nabla \psi_t(r, v) - \sigma(r, v) \psi_t(r, v) \\ &+ \sigma_{\mathbf{s}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{s}}(r, v, v') dv' + \sigma_{\mathbf{f}}(r, v) \int_V \psi_t(r, v') \pi_{\mathbf{f}}(r, v, v') dv', \end{aligned} \quad (4.1)$$

where the five fundamental quantities $\sigma_{\mathbf{s}}$, $\pi_{\mathbf{s}}$, $\sigma_{\mathbf{f}}$, $\pi_{\mathbf{f}}$ and σ (known as *cross-sections* in the physics literature) are all uniformly bounded and measurable with the following interpretation:

- $\sigma_{\mathbf{s}}(r, v)$: the rate at which scattering occurs from incoming velocity v at position r ,
- $\sigma_{\mathbf{f}}(r, v)$: the rate at which fission occurs from incoming velocity v at position r ,
- $\sigma(r, v)$: the sum of the rates $\sigma_{\mathbf{f}} + \sigma_{\mathbf{s}}$ and is known as the total cross section,
- $\pi_{\mathbf{s}}(r, v, v')$: probability density that an incoming velocity v at position r scatters to an outgoing velocity, with probability v' satisfying $\int_V \pi_{\mathbf{s}}(r, v, v') dv' = 1$, and
- $\pi_{\mathbf{f}}(r, v, v')$: density of expected neutron yield at velocity v' from fission with incoming velocity v satisfying $\int_V \pi_{\mathbf{f}}(r, v, v') dv' < \infty$.

It is also usual to assume the additional boundary conditions

$$\begin{cases} \psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \psi_t(r, v) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D, \text{ if } v \cdot \mathbf{n}_r > 0, \end{cases} \quad (4.2)$$

where \mathbf{n}_r is the outward facing normal of D at $r \in \partial D$ and $g : D \times V \rightarrow [0, \infty)$ 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 travelling in the direction of the exterior of the domain are lost to the system.

We will also work with some of (but not necessarily all of) the following assumptions in our results:

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

(H2) We have $\sigma_s \pi_s + \sigma_f \pi_f > 0$ on $D \times V \times V$.

(H3) There is an open ball B compactly embedded in D such that $\sigma_f \pi_f > 0$ on $B \times V \times V$.

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

We note that these assumptions are sufficient but not necessary, and refer the reader to Remark 2.1 in [17] for a discussion of their implications.

4.1.1 Rigorous interpretation of the NTE

As explained in the companion paper [17], the NTE (4.1) is not a meaningful equation in the pointwise sense. Whereas previously (4.1) has been interpreted as an abstract Cauchy process on the $L_2(D \times V)$ space, for probabilistic purposes, the NTE can be better understood in its mild form; see the review discussion in [6]. In particular, the NTE is henceforth understood as the unique bounded solution on bounded intervals of time which satisfy (4.2) and the so-called *mild equation*

$$\psi_t[g](r, v) = \mathbf{U}_t[g](r, v) + \int_0^t \mathbf{U}_s[(\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})\psi_{t-s}[g]](r, v) ds, \quad t \geq 0, r \in D, v \in V. \quad (4.3)$$

for $g \in L_\infty^+(D \times V)$, the space of non-negative functions in $L_\infty(D \times V)$. In (4.3), the advection semigroup is given by

$$\mathbf{U}_t[g](r, v) = g(r + vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)}, \quad t \geq 0. \quad (4.4)$$

where $\kappa_{r,v}^D := \inf\{t > 0 : r + vt \notin D\}$, the scattering operator is given by

$$\overleftarrow{\mathbf{S}}g(r, v) = \sigma_s(r, v) \int_V g(r, v') \pi_s(r, v, v') dv' - \sigma_s(r, v)g(r, v), \quad (4.5)$$

and the fission operator is given by

$$\overleftarrow{\mathbf{F}}g(r, v) = \sigma_f(r, v) \int_V g(r, v') \pi_f(r, v, v') dv' - \sigma_f(r, v)g(r, v), \quad (4.6)$$

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

The papers [17] and [6] discuss in further detail how the mild representation relates to the other classical representation of the NTE via an abstract Cauchy problem which has been treated in e.g. [7, 8, 24]. To understand better why the mild equation (4.3) is indeed a suitable representation for the NTE, we need to understand the probabilistic model of the physical process of nuclear fission.

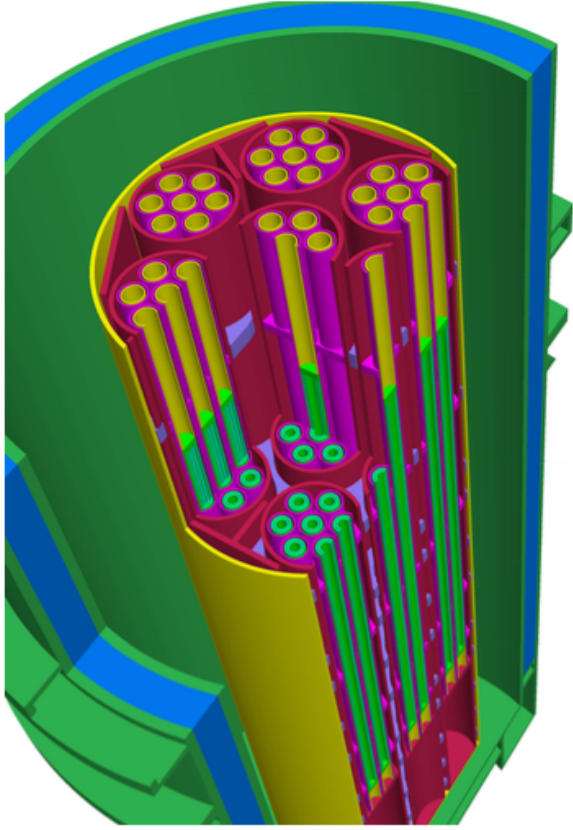


Figure 4-1: The geometry of a nuclear reactor core representing a physical domain D , on to which the different cross-sectional values of $\sigma_s, \sigma_f, \pi_s, \pi_f$ as mapped as numerical values.

4.1.2 Neutron Branching Process

Let us recall from [17], the *neutron branching process* (NBP), whose expectation semigroup provides the solution to (4.3). It is modelled as a branching process, 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 finite atomic measures

$$X_t(A) = \sum_{i=1}^{N_t} \delta_{(r_i(t), v_i(t))}(A), \quad A \in \mathcal{B}(D \times V), \quad t \geq 0, \quad (4.7)$$

where δ is the Dirac measure, 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 measures $\mathcal{M}(D \times V) := \{\sum_{i=1}^n \delta_{(r_i, v_i)} : n \in \mathbb{N}, (r_i, v_i) \in D \times V, i = 1, \dots, n\}$ which evolves randomly as follows.

A particle positioned at r with velocity v will continue to move along the trajectory $r + vt$, until one of the following things happen.

- (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 $T_{\mathbf{s}}$ for the random time that scattering may occur, then providing $r + vt \in D$, independently of the action of fission, $\Pr(T_{\mathbf{s}} > t) = \exp\{-\int_0^t \sigma_{\mathbf{s}}(r + vs, v) ds\}$, for $t \geq 0$.

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

- (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 $T_{\mathbf{f}}$ for the random time that fission may occur, then, providing $r + vt \in D$, independently scattering, $\Pr(T_{\mathbf{f}} > t) = \exp\{-\int_0^t \sigma_{\mathbf{f}}(r + vs, v) ds\}$, for $t \geq 0$.

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 $\{v_i : i = 1, \dots, N\}$. The outgoing velocities are described by the atomic random measure

$$\mathcal{Z}(A) := \sum_{i=1}^N \delta_{v_i}(A), \quad A \in \mathcal{B}(V). \quad (4.8)$$

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 probabilities $\mathcal{P}_{(r,v)}$ are such that, for $v' \in V$, for bounded and measurable $g : V \rightarrow [0, \infty)$,

$$\int_V g(v') \pi_{\mathbf{f}}(r, v, v') dv' = \mathcal{E}_{(r,v)} \left[\int_V g(v') \mathcal{Z}(dv') \right] =: \mathcal{E}_{(r,v)}[\langle g, \mathcal{Z} \rangle], \quad (4.9)$$

where $\mathcal{E}_{(r,v)}$ denotes expectation with respect to $\mathcal{P}_{(r,v)}$. 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).

Write \mathbb{P}_{μ} for the the law of X when issued from an initial configuration $\mu \in \mathcal{M}(D \times V)$. Coming back to how the physical process relates to the NTE, it was show in [6, 17, 7, 8] that, under the assumptions (H1) and (H2), the unique solution, which is bounded on bounded intervals of time, to (4.3) is given by

$$\psi_t[g](r, v) := \mathbb{E}_{\delta_{(r,v)}}[\langle g, X_t \rangle], \quad t \geq 0, r \in \bar{D}, v \in V, \quad (4.10)$$

for $g \in L_{\infty}^+(D \times V)$. The NBP is thus parameterised by the quantities $\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}$ 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 $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \mathcal{P})$ -NBP. It is associated to the NTE via the relation (4.10), but this association does not uniquely identify the NBP. Nonetheless for a given quadruple $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \pi_{\mathbf{f}})$, it was shown in [17] that under the

assumptions (H1) and (H3), at least one NBP always exists that can be associated to it via (4.10).

There is, however, a second equation similar to (4.3), which describes the non-linear semi-group of the neutron branching process and which *does* uniquely identify the $(\sigma_s, \pi_s, \sigma_f, \mathcal{P})$ -NBP. Write the *branching generator* associated with the physical process by⁴

$$G[g](r, v) = \sigma_f(r, v) \mathcal{E}_{(r, v)} \left[\prod_{j=1}^N g(r, v_j) - g(r, v) \right] \quad (4.11)$$

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

$$u_t[g](r, v) := \mathbb{E}_{\delta_{(r, v)}} \left[\prod_{i=1}^{N_t} g(r_i(t), v_i(t)) \right], \quad t \geq 0. \quad (4.12)$$

Formally speaking, by extending the domain in which particles live to include a cemetery state \dagger , corresponding to neutron capture or neutrons going to the boundary ∂D , we will always work with the convention (cf. [19, 20, 21]) that functions appearing in additive functionals are valued as zero on $\{\dagger\}$, whereas in multiplicative functionals, they are valued as one on $\{\dagger\}$. One may think of this as requiring that empty sums are valued as zero where as empty products are valued as one.

As shown in Section 8 of [17], we can break the expectation over the event of scattering or fission in (4.12) and, appealing to standard manipulations (cf. [6, 17]) we see that, for $g \in L_\infty^+(D \times V)$, which is uniformly bounded by unity,

$$u_t[g] = \hat{U}_t[g] + \int_0^t U_s[\bar{S}u_{t-s}[g] + G[u_{t-s}[g]]] ds, \quad t \geq 0, \quad (4.13)$$

where

$$\hat{U}_t[g](r, v) = g(r + v(t \wedge \kappa_{r, v}^D), v). \quad (4.14)$$

Under the assumptions (H1), (H2) and (H4), it was also shown in [17] that (4.13) has a unique solution in the space of non-negative functions, which are bounded over bounded intervals of time.

Before moving on to the asymptotics of $(\psi_t, t \geq 0)$, let us make an important note regarding alternative representations of equations (4.3) and (4.13) for later use. In order to do so, let us momentarily introduce what we mean by a neutron random walk (NRW); cf. [17]. A NRW on D , is defined by its scatter rates, $\alpha(r, v)$, $r \in D, v \in V$, and scatter probability densities $\pi(r, v, v')$, $r \in D, v, v' \in V$ where $\int_V \pi(r, v, v') dv' = 1$ for all $r \in D, v \in V$. When issued from $r \in D$ 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_s a scattering occurs, where $\Pr(T_s > t) = \exp\{-\int_0^t \alpha(r + vs, v) ds\}$, for $t \geq 0$. When the scattering event occurs in

⁴Here and elsewhere, an empty product is always understood to be unity.

position-velocity configuration (r, v) , a new velocity v' is selected with probability $\pi(r, v, v')dv'$. We refer more specifically to the latter as an $\alpha\pi$ -NRW.

The linear mild equation (4.3) and its accompanying non-linear mild form (4.13), although consistent with existing literature (cf. [6, 17, 18, 5]) can be equally identified as the unique (in the same sense as mentioned in the previous paragraph) solution to the equations

$$\psi_t[g](r, v) = \mathbf{Q}_t[g](r, v) + \int_0^t \mathbf{Q}_s[\tilde{\mathbf{F}}\psi_{t-s}[g]](r, v)ds, \quad t \geq 0, r \in D, v \in V. \quad (4.15)$$

and

$$u_t[g] = \hat{\mathbf{Q}}_t[g](r, v) + \int_0^t \mathbf{Q}_s[G[u_{t-s}[g]]](r, v)ds, \quad t \geq 0, r \in D, v \in V, \quad (4.16)$$

respectively, where for $g \in L_\infty^+(D \times V)$,

$$\mathbf{Q}_t[g](r, v) = \mathbf{E}_{(r,v)}[f(R_t, \Upsilon_t)\mathbf{1}_{(t < \tau^D)}],$$

and

$$\hat{\mathbf{Q}}_t[g](r, v) = \mathbf{E}_{(r,v)}[f(R_{t \wedge \tau^D}, \Upsilon_{t \wedge \tau^D})],$$

are the expectation semigroups associated with the $\sigma_{\mathfrak{s}}\pi_{\mathfrak{s}}$ -NRW and $\tau^D = \inf\{t > 0 : R_t \notin D\}$.

4.1.3 Lead order asymptotics of the expectation semigroup

In the accompanying predecessor to this article, [17], a Perron-Frobenius type asymptotic was developed for $(\psi_t, t \geq 0)$. In order to state it we need to introduce another assumption, which is slightly stronger than (H2). To this end, define

$$\alpha(r, v)\pi(r, v, v') = \sigma_{\mathfrak{s}}(r, v)\pi_{\mathfrak{s}}(r, v, v') + \sigma_{\mathfrak{f}}(r, v)\pi_{\mathfrak{f}}(r, v, v') \quad r \in D, v, v' \in V. \quad (4.17)$$

Our new condition is:

$$\text{(H2)*: We have } \inf_{r \in D, v, v' \in V} \alpha(r, v)\pi(r, v, v') > 0.$$

Theorem 11. *Suppose that (H1) and (H2)* hold. Then, for semigroup $(\psi_t, t \geq 0)$ identified by (4.3), there exists a $\lambda_* \in \mathbb{R}$, a positive⁵ right eigenfunction $\varphi \in L_\infty^+(D \times V)$ and a left eigenmeasure which is absolutely continuous with respect to Lebesgue measure on $D \times V$ with density $\tilde{\varphi} \in L_\infty^+(D \times V)$, both having associated eigenvalue $e^{\lambda_* t}$, and such that φ (resp. $\tilde{\varphi}$) is uniformly (resp. a.e. uniformly) bounded away from zero on each compactly embedded subset of $D \times V$. In particular for all $g \in L_\infty^+(D \times V)$*

$$\langle \tilde{\varphi}, \psi_t[g] \rangle = e^{\lambda_* t} \langle \tilde{\varphi}, g \rangle \quad (\text{resp. } \psi_t[\varphi] = e^{\lambda_* t} \varphi) \quad t \geq 0. \quad (4.18)$$

⁵To be precise, by a positive eigenfunction, we mean a mapping from $D \times V \rightarrow (0, \infty)$. This does not prevent it being valued zero on ∂D , as D is an open bounded, convex domain.

Moreover, there exists $\varepsilon > 0$ such that, for all $g \in L_\infty^+(D \times V)$,

$$\left\| e^{-\lambda_* t} \varphi^{-1} \psi_t[g] - \langle \tilde{\varphi}, g \rangle \right\|_\infty = O(e^{-\varepsilon t}) \text{ as } t \rightarrow +\infty. \quad (4.19)$$

In light of Theorem 11, we can categorise the physical process according to the value of λ_* . In particular, when $\lambda_* > 0$ we say the process is supercritical, when $\lambda_* = 0$, the process is critical and when $\lambda_* < 0$, the process is subcritical.

4.1.4 Strong law of large numbers at supercriticality

The main aim of this article as a continuation of [17] is to understand the almost sure behaviour of the $(\sigma_s, \pi_s, \sigma_f, \mathcal{P})$ -NBP in relation to what is, in effect, a statement of mean growth in Theorem 11, in the setting that $\lambda_* > 0$. In the aforesaid article, it was noted that

$$W_t := e^{-\lambda_* t} \frac{\langle \varphi, X_t \rangle}{\langle \varphi, \mu \rangle}, \quad t \geq 0, \quad (4.20)$$

is a unit mean martingale under \mathbb{P}_μ , $\mu \in \mathcal{M}(D \times V)$ and, moreover its convergence was studied. In particular, since the martingale is non-negative, we automatically know that it must converge to a limiting random variable, that is, $W_t \rightarrow W_\infty$, \mathbb{P}_μ -almost surely, where we can take $W_\infty := \liminf_{t \rightarrow \infty} W_t$ for definiteness. Before stating the result regarding the latter, we require one more assumption on the NBP:

(H3)*: There exists a ball B compactly embedded in D such that

$$\inf_{r \in B, v, v' \in V} \sigma_f(r, v) \pi_f(r, v, v') > 0.$$

Theorem 12 ([17]). *For the $(\sigma_s, \pi_s, \sigma_f, \mathcal{P})$ -NBP satisfying the assumptions (H1), (H2)*, (H3)* and (H4), the martingale $(W_t, t \geq 0)$ converges to W_∞ \mathbb{P} -almost surely and in $L_2(\mathbb{P})$ if and only if $\lambda_* > 0$, otherwise $W_\infty = 0$ \mathbb{P} -almost surely.*

Note that when $\lambda_* \leq 0$, since $\lim_{t \rightarrow \infty} W_t = 0$ almost surely, it follows that, for each Ω compactly embedded in $D \times V$, $\lim_{t \rightarrow \infty} X_t(\Omega) = 0$. It therefore remains to describe the growth of $X_t(\Omega)$, $t \geq 0$, for $\lambda_* > 0$. This is the main result of this paper, given below. In order to state it, we must introduce the notion of a directionally continuous function on $D \times V$. Such functions are defined as having the property that, for all $r \in D$, $v \in V$,

$$\lim_{t \downarrow 0} g(r + vt, v) = g(r, v).$$

Theorem 13. *Suppose the assumptions of Theorem 12 hold. For all measurable and directionally continuous non-negative g on $D \times V$ such that, up to a multiplicative constant, $g \leq \varphi$, then for*

any initial configuration $\mu \in \mathcal{M}(D \times V)$,

$$e^{-\lambda_* t} \frac{\langle g, X_t \rangle}{\langle \varphi, \mu \rangle} \rightarrow \langle g, \tilde{\varphi} \rangle W_\infty$$

\mathbb{P}_μ -almost surely and in $L_2(\mathbb{P})$, as $t \rightarrow \infty$.

To the best of our knowledge no such results can be found in the existing neutron transport literature. The closest known results are found in the final section of [25] and are significantly weaker than Theorem 13.

We can think of Theorem 13 as stating a stochastic analogue of (4.19), noting, for example, that the former implies

$$\lim_{t \rightarrow \infty} e^{-\lambda_* t} \frac{\mathbb{E}_{\delta(r,v)}[\langle g, X_t \rangle]}{\varphi(r,v)} = \langle g, \tilde{\varphi} \rangle \quad (4.21)$$

for all $r \in D, v \in V$, which is a version of the latter (albeit without the speed of convergence).

The proof of Theorem 13 relies on a fundamental path decomposition, often referred to in the theory of spatial and non-spatial branching processes as a *skeletal decomposition*, see e.g. [12, 14, 2, 9, 26]. The skeletal decomposition is essential in that it identifies an embedded NBP within the original one for which there is no neutron-absorption (neither at ∂D nor into nuclei at collision). This ‘thinned down tree’ is significantly easier to analyse for technical reasons, but nonetheless provides all the mass in the limit (4.21).

4.2 Skeletal decomposition

Inspired by [14], we dedicate this section to the proof of a so-called skeletal decomposition, which necessarily requires us to have $\lambda_* > 0$. In very rough terms, for the NBP, we can speak of genealogical lines of descent, meaning neutrons that came from a fission event of a neutron that came from a fission event of a neutron ... and so on, back to one of the initial neutrons a time $t = 0$. If we focus on an individual genealogical line of descent embedded in the NBP, it has a space-velocity trajectory which takes the form of a NRW whose spatial component may or may not hit the boundary of D . Indeed, when the NBP survives for all time (requiring $\lambda_* > 0$), there must necessarily be some genealogical lines of descent whose spatial trajectories remain in D forever.

The basic idea of the skeletal decomposition is to consider the collection of all surviving genealogical lines of descent and understand their space-velocity dynamics collectively as a process (*the skeleton*). It turns out that the skeleton forms another NBP but with different scatter and fission statistics from the underlying NBP, due to the fact that we are considering genealogical lines of descent which are biased, since they remain in D for all time. For the remaining neutron trajectories that go to the boundary of D or end in neutron capture, the skeletal decomposition identifies them as immigrants that are thrown off the path of the skeleton.

Below, we develop the statement of the skeletal decomposition. It was brought to our attention by a referee that the proof is robust enough to work in the relatively general setting of a *Markov branching process (MBP) with non-local branching* and hence we first set up the notation of a general branching process. It is worthy of note that the motivation for this switch to a general setting is that, for branching particle systems, nothing is known of skeletal decompositions for non-local branching generators; although some results have been identified in the more continuous setting of superprocesses, cf [26], they do not apply to particle systems. Our proof is inspired by the martingale arguments found in [14] which gives a skeletal decomposition for branching Brownian motion in a strip with local branching.

4.2.1 The general branching Markov setup

Until the end of this section (Section 4.2), unless otherwise mentioned, we will work in the setting of a general MBP, which we will shortly define in more detail. The reader will note that we necessarily choose to overlap our notation for this general setting with that of the NBP. As such, the reader is encouraged to keep in mind the application to the NBP at all times. Additionally, we provide some remarks at the end of this section to illustrate how the general case takes a specific form in the case of the NBP.

Henceforth, $X = (X_t, t \geq 0)$ will be a (\mathbf{P}, G) -Markov branching process on a non-empty, open Euclidian domain⁶ $E \subseteq \mathbb{R}^d$, where $\mathbf{P} = (\mathbf{P}_t, t \geq 0)$ is a Markov semigroup on E and G is the associated branching generator. More precisely, X is an atomic measure-valued stochastic process (in a similar sense to (4.7)) in which particles move independently according to a copy of the Markov process associated to \mathbf{P} such that, when a particle is positioned at $x \in E$, at the instantaneous spatial rate $\varsigma(x)$, the process will branch and a random number of offspring, say N , are thrown out in positions, say x_1, \dots, x_N in E , according to some law \mathcal{P}_x . (Note, we always consider of (x_1, \dots, x_N) as an ordered set of points.)

We do not need \mathbf{P} to have the Feller property, and we assume nothing of the boundary conditions on E , in particular, \mathbf{P} need not be conservative. That said, it will prove to be more convenient to introduce a (possible) cemetery state \dagger appended to E , which is to be treated as an absorbing state, and regard \mathbf{P} as conservative. As such,

$$\mathbf{P}_t[f](x) = \mathbf{E}_x[f(\xi_t)] = \mathbf{E}_x[f(\xi_t)\mathbf{1}_{(t < \mathbf{k})}], \quad x \in E, f \in L_\infty^+(E), \quad (4.22)$$

where the process ξ , with probabilities $(\mathbf{P}_x, x \in E)$, is the Markov process on $E \cup \{\dagger\}$ with lifetime $\mathbf{k} = \inf\{t > 0 : \xi_t \in \{\dagger\}\}$, $L_\infty^+(E)$ is the space of bounded, measurable functions on E and, in this context, we always take $f(\dagger) := 0$.

As such, in a similar spirit to (4.11), we can think of the branching generator, G , as having definition

$$G[f](x) = \varsigma(x)\mathcal{E}_x \left[\prod_{j=1}^N f(x_j) - f(x) \right], \quad x \in E, \quad (4.23)$$

⁶The arguments presented here are robust enough to work with more abstract domains; see for example the set up in [1].

for $f \in L_\infty^{+,1}(E)$, the space of non-negative measurable functions on E bounded by unity. As previously, we always define the empty product as equal to unity.

We use \mathbb{P}_{δ_x} for the law of X issued from a single particle positioned at $x \in E$. In a similar spirit to (4.12), we can introduce the non-linear semigroup of the branching process,

$$u_t[g](x) := \mathbb{E}_{\delta_x} \left[\prod_{i=1}^{N_t} g(x_i(t)) \right], \quad t \geq 0, x \in E, g \in L_\infty^{+,1}(E), \quad (4.24)$$

where $X_t = \sum_{i=1}^{N_t} \delta_{x_i(t)}$, $t \geq 0$. As before, we define the empty product to be unity, and for consistency, functions, g , appearing in such functionals can be valued on $E \cup \{\dagger\}$ and forced to take the value $g(\dagger) = 1$.

Similarly to the derivation of (4.13) and (4.16), it is straightforward to show that, for such functions, $u_t[g]$ solves the non-linear mild equation

$$u_t[g] = \hat{\mathbb{P}}_t[g](x) + \int_0^t \mathbb{P}_s[G[u_{t-s}[g]]](x) ds, \quad t \geq 0, x \in E, \quad (4.25)$$

where we need to adjust \mathbb{P} to $\hat{\mathbb{P}}$ to accommodate for the fact that empty products are valued as one, as follows

$$\hat{\mathbb{P}}_t[g](x) = \mathbf{E}_x[g(\xi_{t \wedge \mathbf{k}})], \quad x \in E. \quad (4.26)$$

Now, define

$$\zeta := \inf\{t \geq 0 : \langle 1, X_t \rangle = 0\}, \quad (4.27)$$

the time of extinction, and let

$$w(x) := \mathbb{P}_{\delta_x}(\zeta < \infty). \quad (4.28)$$

We will also frequently use with

$$p(x) := 1 - w(x), \quad x \in E.$$

Recalling that we need to take as a definition $w(\dagger) = 1$, by conditioning on $\mathcal{F}_t = \sigma(X_s, s \leq t)$, for $t \geq 0$,

$$w(x) = \mathbb{E}_{\delta_x} \left[\prod_{i=1}^{N_t} w(x_i(t)) \right]. \quad (4.29)$$

Taking (4.29), (4.25) and (4.26) into account, it is easy to deduce that w also solves

$$w(x) = \hat{\mathbb{P}}_t[w](x) + \int_0^t \mathbb{P}_s[G[w]](x) ds, \quad t \geq 0, x \in E. \quad (4.30)$$

We will assume:

(M1): $\inf_{x \in E} w(x) > 0$ and $w(x) < 1$ for $x \in E$.

Beyond this, we assume relatively little about \mathbb{P} and G other than:

(M2): The branching rate ς is uniformly bounded from above.

Re-writing (4.30) in the form

$$w(x) = \mathbf{E}_x[w(\xi_{t \wedge k})] + \mathbf{E}_x \left[\int_0^{t \wedge k} w(\xi_s) \frac{G[w](\xi_s)}{w(\xi_s)} ds \right], \quad t \geq 0,$$

and noting that $\sup_{x \in E} G[w](x)/w(x) < \infty$, we can appeal to the method of exchanging exponential potential for additive potential⁷ in e.g. [10, Lemma 1.2, Chapter 4, Part 1], which yields

$$w(x) = \mathbf{E}_x \left[w(\xi_{t \wedge k}) \exp \left(\int_0^{t \wedge k} \frac{G[w](\xi_s)}{w(\xi_s)} ds \right) \right], \quad x \in E, t \geq 0. \quad (4.31)$$

This identity will turn out to be extremely useful in our analysis, in particular, the equality (4.31) together with the Markov property of ξ implies that the object in the expectation on the right-hand side of (4.31) is a martingale.

In Theorem 14 below we give the skeletal decomposition in the form of a theorem. In order to state this result, we first need to develop two notions of conditioning. As there is rather a lot of notation, we include a table in the Appendix which the reader may refer to as needed.

The basic pretext of the skeletal decomposition is that we want to split genealogical lines of descent into those that survive forever and those that reach a dead end. To this end, let $c_i(t)$ denote the label of a particle $i \in \{1, \dots, N_t\}$. We label a particle ‘prolific’, denoted $c_i(t) = \uparrow$, if it has an infinite genealogical line of descent, and $c_i(t) = \downarrow$, if its line of descent dies out (i.e. ‘non-prolific’). Ultimately, we want to describe how the spatial genealogical tree of the MBP can be split into into a spatial genealogical sub-tree, consisting of \uparrow -labelled particles (the skeleton), which is dressed with trees of \downarrow -labelled particles.

Let $\mathbb{P}^\updownarrow = (\mathbb{P}_{\delta_x}^\updownarrow, x \in E)$ denote the probabilities of the two-labelled process described above. Then for $t \geq 0$ and $x \in E$ we have the following relationship between \mathbb{P}^\updownarrow and \mathbb{P} :

$$\frac{d\mathbb{P}_{\delta_x}^\updownarrow}{d\mathbb{P}_{\delta_x}} \Big|_{\mathcal{F}_\infty} = \prod_{i=1}^{N_t} (\mathbf{1}_{(c_i(t)=\uparrow)} + \mathbf{1}_{(c_i(t)=\downarrow)}) = 1, \quad (4.32)$$

where $\mathcal{F}_\infty = \sigma(\cup_{t \geq 0} \mathcal{F}_t)$. Projecting onto \mathcal{F}_t , for $t \geq 0$, we have

$$\begin{aligned} \frac{d\mathbb{P}_{\delta_x}^\updownarrow}{d\mathbb{P}_{\delta_x}} \Big|_{\mathcal{F}_t} &= \mathbb{E}_{\delta_x} \left(\prod_{i=1}^{N_t} (\mathbf{1}_{(c_i(t)=\uparrow)} + \mathbf{1}_{(c_i(t)=\downarrow)}) \Big| \mathcal{F}_t \right) \\ &= \sum_{I \subseteq \{1, \dots, N_t\}} \prod_{i \in I} \mathbb{P}_{\delta_x}(c_i(t) = \uparrow | \mathcal{F}_t) \prod_{i \in \{1, \dots, N_t\} \setminus I} \mathbb{P}_{\delta_x}(c_i(t) = \downarrow | \mathcal{F}_t) \\ &= \sum_{I \subseteq \{1, \dots, N_t\}} \prod_{i \in I} p(x_i(t)) \prod_{i \in \{1, \dots, N_t\} \setminus I} w(x_i(t)), \end{aligned} \quad (4.33)$$

⁷We will use this trick throughout this paper and consistently refer to it as the ‘transfer of the exponential potential to the additive potential’ and vice versa in the other direction.

where we understand the sum to be taken over all subsets of $\{1, \dots, N_t\}$, each of which is denoted by I .

The decomposition in (4.33) indicates the beginning point of how we break up the law of the (P, G) -MBP according to subtrees that are categorised as \downarrow (with probability w) and subtrees that are categorised as \uparrow with \downarrow dressing (with probability p), the so-called *skeletal decomposition*.

In the next two sections we will examine the notion of our MBP conditioned to die out and conditioned to survive, respectively. Thereafter we will use the characterisation of these conditioned trees to formalise our skeletal decomposition.

4.2.2 \downarrow -trees

Following [14], let us start by characterising the law of genealogical trees populated by the marks \downarrow . Thanks to the branching property, it suffices to consider trees which are issued with a single particle with mark \downarrow . By definition of the mark $c_\emptyset(0) = \downarrow$, where \emptyset is the initial ancestral particle, this is the same as understanding the law of (X, \mathbb{P}) conditioned to become extinct. Indeed, for $A \in \mathcal{F}_t$,

$$\begin{aligned} \mathbb{P}_{\delta_x}^\downarrow(A) &:= \mathbb{P}_{\delta_x}^\uparrow(A | c_\emptyset(0) = \downarrow) \\ &= \frac{\mathbb{P}_{\delta_x}^\uparrow(A; c_i = \downarrow, \text{ for each } i = 1, \dots, N_t)}{\mathbb{P}_{\delta_x}^\uparrow(c_\emptyset(0) = \downarrow)} \\ &= \frac{\mathbb{E}_{\delta_x} \left[\mathbf{1}_A \prod_{i=1}^{N_t} w(x_i(t)) \right]}{w(x)}. \end{aligned} \quad (4.34)$$

We are now in a position to characterise the MBP trees which are conditioned to become extinct (equivalently, with genealogical lines of descent which are marked entirely with \downarrow). Heuristically speaking, the next proposition shows that the conditioning creates a branching particle process in which particles are prone to die out (whether that be due to being killed at the boundary under P , or suppressing offspring). Our proof is partly inspired by Proposition 11 of [14].

Proposition 3 (\downarrow Trees). *For initial configurations of the form $\nu = \sum_{i=1}^n \delta_{x_i}$, for $n \in \mathbb{N}$ and $x_1, \dots, x_n \in E$, define the measure $\mathbb{P}_\nu^\downarrow$ via (4.34),*

$$\mathbb{P}_\nu^\downarrow = \otimes_{i=1}^n \mathbb{P}_{\delta_{x_i}}^\downarrow,$$

i.e. starting independent processes at positions x_i each under $\mathbb{P}_{\delta_{x_i}}$, for $i = 1, \dots, n$. Then under $\mathbb{P}_\nu^\downarrow$, X is a $(P^\downarrow, G^\downarrow)$ -MBP with motions semigroup P^\downarrow and branching generator G^\downarrow defined as follows. The motion semigroup P^\downarrow is that of the Markov process ξ with probabilities $(\mathbf{P}_x^\downarrow, x \in E)$, where

$$\left. \frac{d\mathbf{P}_x^\downarrow}{d\mathbf{P}_x} \right|_{\sigma(\xi_s, s \leq t)} = \frac{w(\xi_{t \wedge k})}{w(x)} \exp \left(\int_0^{t \wedge k} \frac{G[w](\xi_s)}{w(\xi_s)} ds \right), \quad t \geq 0. \quad (4.35)$$

For $x \in E$ and $f \in L_\infty^{+,1}(E)$, the branching generator is given by

$$G^\downarrow[f] = \varsigma^\downarrow(x) \mathcal{E}_x^\downarrow \left[\prod_{j=1}^N f(x_j) - f(x) \right] = \frac{1}{w} [G[fw] - fG[w]], \quad (4.36)$$

where,

$$\varsigma^\downarrow(x) = \varsigma(x) + \frac{G[w](x)}{w(x)} = \frac{\varsigma(x)}{w(x)} \mathcal{E}_x \left[\prod_{j=1}^N w(x_j) \right] \quad x \in E, \quad (4.37)$$

and

$$\frac{d\mathcal{P}_x^\downarrow}{d\mathcal{P}_x} \Big|_{\sigma(N, x_1, \dots, x_N)} = \frac{\prod_{i=1}^N w(x_i)}{\mathcal{E}_x \left[\prod_{j=1}^N w(x_j) \right]} = \frac{\varsigma(x)}{\varsigma^\downarrow(x)w(x)} \prod_{i=1}^N w(x_i) \quad (4.38)$$

Proof of Proposition 3. First let us show that the change of measure results in a particle process that respects the Markov branching property. In a more general sense, for ν as in the statement of this proposition, (4.34) takes the form

$$\frac{d\mathbb{P}_\nu^\downarrow}{d\mathbb{P}_\nu} \Big|_{\mathcal{F}_t} = \frac{\prod_{i=1}^{N_t} w(x_i(t))}{\prod_{i=1}^{n_t} w(x_i)}$$

It is clear from the conditioning that every particle in the resulting process under the new measure $\mathbb{P}_\nu^\downarrow$ must carry the mark \downarrow , i.e. be non-prolific, by construction.

Let us define, for $g \in L_\infty^{+,1}(E)$,

$$u_t^\downarrow[g](x) = \mathbb{E}_{\delta_x}^\downarrow \left[\prod_{i=1}^{N_t} g(x_i(t)) \mid c_\emptyset(0) = \downarrow \right] = \frac{1}{w(x)} u_t[wg](x), \quad (4.39)$$

which describes the evolution of the the process X under \mathbb{P}^\downarrow . In particular, for $g \in L_\infty^{+,1}(E)$, $x \in E$ and $s, t \geq 0$, note that

$$\begin{aligned} \mathbb{E}_{\delta_x}^\downarrow \left[\prod_{i=1}^{N_{t+s}} g(x_i(t+s)) \mid \mathcal{F}_t \right] &= \frac{1}{w(x)} \prod_{i=1}^{N_t} w(x_i(t)) \mathbb{E}_{\delta_x} \left[\frac{\prod_{j=1}^{N_s^i} w(x_j^i(s)) g(x_j^i(s))}{w(x_i(t))} \mid \mathcal{F}_t \right] \\ &= \frac{1}{w(x)} \prod_{i=1}^{N_t} w(x_i(t)) u_s^\downarrow[g](x_i(t)), \end{aligned} \quad (4.40)$$

where, given \mathcal{F}_t , $((x_j^i(t)), j = 1, \dots, N_s^i)$ are the physical configurations of particles at time $t+s$ that are descendent from particle $i \in N_t$. This ensures the Markov branching property holds.

It thus suffices for the remainder of the proof to show, in the spirit of (4.13), that, for $g \in L_\infty^{+,1}(E)$,

$$u_t^\downarrow[g](x) = \hat{\mathbb{P}}_t^\downarrow[g](x) + \int_0^t \mathbb{P}_s^\downarrow[G^\downarrow[u_{t-s}^\downarrow[g]]](x) ds, \quad t \geq 0, x \in E. \quad (4.41)$$

holds, where $\hat{\mathbb{P}}^\downarrow$ is defined in a similar spirit to (4.26), which is the semigroup evolution equation

for a $(\mathbb{P}^\downarrow, G^\downarrow)$ -MBP, and to identify the internal structure of G^\downarrow .

From (4.25) and (4.39) it follows that, for $g \in L_\infty^{+,1}(E)$,

$$u_t^\downarrow[g] = \frac{1}{w} \hat{\mathbb{P}}_t[wg] + \int_0^t \frac{1}{w} \mathbb{P}_s[G[wu_{t-s}^\downarrow[g]]] ds, \quad t \geq 0, \quad (4.42)$$

In the spirit of the derivation of (4.31), we can apply [11, Lemma 1.2, Chapter 4, Part 1] and use (4.35) and (4.42) to get, for $x \in E$,

$$\begin{aligned} u_t^\downarrow[g](x) &= \frac{1}{w(x)} \hat{\mathbb{P}}_t[wg](x) + \int_0^t \frac{1}{w(x)} \mathbb{P}_s \left[w \frac{G[wu_{t-s}^\downarrow[g]]}{w} \right] (x) ds \\ &+ \int_0^t \frac{1}{w(x)} \mathbb{P}_s \left[\frac{G[w]}{w} wu_{t-s}^\downarrow[g] \right] (x) ds - \int_0^t \frac{1}{w(x)} \mathbb{P}_s \left[\frac{G[w]}{w} wu_{t-s}^\downarrow[g] \right] (x) ds \\ &= \frac{1}{w(x)} \mathbf{E}_x \left[g(\xi_{t \wedge k}) w(\xi_{t \wedge k}) e^{\int_0^{t \wedge k} \frac{G[w](\xi_u)}{w(\xi_u)} du} \right] \\ &+ \frac{1}{w(x)} \mathbf{E}_x \left[\int_0^{t \wedge k} \frac{G[wu_{t-s}^\downarrow[g]](\xi_s)}{w(\xi_s)} w(\xi_s) e^{\int_0^s \frac{G[w](\xi_u)}{w(\xi_u)} du} ds \right] \\ &- \frac{1}{w(x)} \mathbf{E}_x \left[\int_0^{t \wedge k} \frac{G[w](\xi_s)}{w(\xi_s)} u_{t-s}^\downarrow[g](\xi_s) w(\xi_s) e^{\int_0^s \frac{G[w](\xi_u)}{w(\xi_u)} du} ds \right] \\ &= \hat{\mathbb{P}}_t^\downarrow[g](x) + \int_0^t \mathbb{P}_s^\downarrow \left[\frac{G[wu_{t-s}^\downarrow[g]]}{w} \right] (x) ds - \int_0^t \mathbb{P}_s^\downarrow \left[\frac{G[w]}{w} u_{t-s}^\downarrow[g] \right] (x) ds \\ &= \hat{\mathbb{P}}_t^\downarrow[g](x) + \int_0^t \mathbb{P}_s^\downarrow \left[G^\downarrow[u_{t-s}^\downarrow[g]] \right] (x) ds \end{aligned}$$

where we have used the definition (4.36).

It remains to identify the internal structure of G^\downarrow . Recalling that $\varsigma^\downarrow = \varsigma + w^{-1}G[w]$, we have, for $f \in L_\infty^{+,1}(E)$,

$$\begin{aligned} G^\downarrow[f](x) &= \frac{1}{w(x)} [G[fw] - fG[w]](x) \\ &= \frac{1}{w(x)} \left[\varsigma(x) \mathcal{E}_x \left[\prod_{i=1}^N f(x_i) w(x_i) \right] - \varsigma(x) f(x) w(x) - fG[w](x) \right] \\ &= \frac{\varsigma(x)}{w(x)} \mathcal{E}_x \left[\prod_{i=1}^N f(x_i) w(x_i) \right] - \left(\varsigma(x) + \frac{G[w](x)}{w} \right) f(x) \\ &= \varsigma^\downarrow(x) \left(\frac{\varsigma(x)}{\varsigma^\downarrow(x) w(x)} \mathcal{E}_x \left[\prod_{i=1}^N w(x_i) f(x_i) \right] - f(x) \right), \end{aligned}$$

Moreover, recalling the change of measure (4.38), note that, for $x \in E$, \mathbb{P}_x^\downarrow is a probability measure on account of the fact that, when we set $f \equiv 1$, recalling again that $\varsigma^\downarrow = \varsigma + w^{-1}G[w]$

as well as the definition of G given in (4.23),

$$\mathcal{E}_x \left[\frac{\varsigma(x)}{\varsigma^\downarrow(x)w(x)} \prod_{i=1}^N w(x_i) \right] = \frac{G[w](x) + \varsigma(x)w(x)}{\varsigma(x) + w^{-1}(x)G[w](x)} \frac{1}{w(x)} = 1$$

as required. \square

4.2.3 \updownarrow -trees

In a similar spirit to the previous section we can look at the law of our MBP, when issued from a single ancestor, conditioned to have a subtree of prolific individuals. As such, for $A \in \mathcal{F}_t$, we define

$$\begin{aligned} \mathbb{P}_{\delta_x}^{\updownarrow}(A | c_\emptyset(0) = \uparrow) &= \frac{\mathbb{P}_{\delta_x}^{\updownarrow}(A; c_i = \uparrow, \text{ for at least one } i = 1, \dots, N_t)}{\mathbb{P}_{\delta_x}^{\updownarrow}(c_\emptyset(0) = \uparrow)} \\ &= \frac{\mathbb{E}_{\delta_x} \left[\mathbf{1}_A \left(1 - \prod_{i=1}^{N_t} w(x_i(t)) \right) \right]}{p(x)}. \end{aligned} \quad (4.43)$$

In the next proposition, we want to describe our MBP under $\mathbb{P}_{\delta_x}^{\updownarrow}(\cdot | c_\emptyset(0) = \uparrow)$. In order to do so, we first need to introduce a type- \uparrow -type- \downarrow MBP.

Our type- \uparrow -type- \downarrow MBP process, say $X^\updownarrow = (X_t^\updownarrow, t \geq 0)$, has an ancestor which is of type- \uparrow . We will implicitly assume (and suppress from the notation X^\updownarrow) that $X_0^\updownarrow = \delta_x$, for $x \in E$. Particles in X^\updownarrow of type- \uparrow move as a \mathbf{P}^\uparrow -Markov process. When a branching event occurs for a type- \uparrow particle, both type- \uparrow and type- \downarrow particles may be produced, but always at least one type- \uparrow is produced. Type- \uparrow particles may be thought of as offspring and any additional type- \downarrow particles may be thought of as immigrants. Type- \downarrow particles that are created can only subsequently produce type- \downarrow particles in such a way that they give rise to a $(\mathbf{P}^\downarrow, G^\downarrow)$ -MBP.

The joint branching/immigration rate of type- \uparrow and type- \downarrow particles in X^\updownarrow at $x \in E$ is given by

$$\varsigma^\updownarrow(x) = \frac{\varsigma(x)}{p(x)} \mathcal{E}_x \left[1 - \prod_{j=1}^N w(x_j) \right]. \quad (4.44)$$

We can think of the branching rate in (4.44) as the original rate $\varsigma(x)$ multiplied by the probability (under \mathcal{P}_x) that at least one of the offspring are of type- \uparrow , given the branching particle is of type- \uparrow

At a branching/immigration event of a type- \uparrow particle, we will write N^\uparrow and $(x_i^\uparrow, i = 1, \dots, N^\uparrow)$ for the number and positions of type- \uparrow offspring and N^\downarrow and $(x_j^\downarrow, j = 1, \dots, N^\downarrow)$ for the number and positions of type- \downarrow immigrants. We will write $(\mathcal{P}_x^\updownarrow, x \in E)$ for the joint law of the the random variables in the previous sentence. Formally speaking, the branching generator,

G^\uparrow , of offspring/immigrants for a type- \uparrow particle positioned at $x \in E$ is written

$$G^\uparrow[f, g](x) = \varsigma^\uparrow(x) \left(\mathcal{E}_x^\uparrow \left[\prod_{i=1}^{N^\uparrow} f(x_i^\uparrow) \prod_{j=1}^{N^\downarrow} g(x_j^\downarrow) \right] - f(x) \right) \quad (4.45)$$

for $f, g \in L_\infty^{+,1}(E)$.

For our process X^\uparrow , for each $x \in E$, we will define the laws \mathcal{P}_x^\uparrow in terms of an additional random selection from $(x_i, i = 1, \dots, N)$ under \mathcal{P}_x . Write \mathcal{N}^\uparrow for the set of indices in $\{1, \dots, N\}$, which, together, identify the type- \uparrow particles, i.e. $(x_i, i \in \mathcal{N}^\uparrow) = (x_j^\uparrow, j = 1, \dots, N^\uparrow)$. The remaining indices $\{1, \dots, N\} \setminus \mathcal{N}^\uparrow$ will identify the type- \downarrow immigrants from $(x_i, i = 1, \dots, N)$. Thus, to describe \mathcal{P}_x^\uparrow , for any $x \in E$, it suffices to give the law of $(N; x_1, \dots, x_N; \mathcal{N}^\uparrow)$. To this end, for $F \in \sigma(N; x_1, \dots, x_N)$ and $I \subseteq \mathbb{N}$, we will set

$$\mathcal{P}_x^\uparrow(F \cap \{\mathcal{N}^\uparrow = I\}) := \mathbf{1}_{\{|I| \geq 1\}} \frac{\varsigma(x)}{\varsigma^\uparrow(x)p(x)} \mathcal{E}_x \left[\mathbf{1}_{F \cap \{I \subseteq \{1, \dots, N\}\}} \prod_{i \in I} p(x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i) \right]. \quad (4.46)$$

Equivalently, we could define $\mathcal{P}_x^\uparrow(F) := \mathcal{P}_x(F)$ for all $F \in \sigma(N; x_1, \dots, x_N)$, and, for all $I \subseteq \mathbb{N}$,

$$\mathcal{P}_x^\uparrow(\mathcal{N}^\uparrow = I | \sigma(N; x_1, \dots, x_N)) := \mathbf{1}_{\{|I| \geq 1\} \cap \{I \subseteq \{1, \dots, N\}\}} \frac{\prod_{i \in I} p(x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i)}{1 - \mathcal{E}_x \left[\prod_{j=1}^N w(x_j) \right]}. \quad (4.47)$$

The pairs $(x_i^\uparrow, i = 1, \dots, N^\uparrow)$ and $(x_j^\downarrow, j = 1, \dots, N^\downarrow)$ under $(\mathcal{P}_x^\uparrow, x \in E)$ in (4.45) can thus be seen as equal in law to selecting the type of each particle following an independent sample of the non-local branching configuration (x_1, \dots, x_N) under \mathcal{P}_x , where each x_k is independently assigned either as type- \uparrow with probability $p(x_k)$ or as type- \downarrow with probability $w(x_k) = 1 - p(x_k)$, but then conditional on there being at least one type- \uparrow .

As such with the definitions above, it is now a straightforward exercise to identify the branching generator in (4.45) in terms of $(x_i, i = 1, \dots, N)$ under $(\mathcal{P}_x, x \in E)$ via the following identity, for $x \in E$,

$$G^\uparrow[f, g](x) = \frac{\varsigma(x)}{p(x)} \mathcal{E}_x \left[\sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \prod_{i \in I} p(x_i) f(x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i) g(x_i) \right] - \varsigma^\uparrow(x) f(x) \quad (4.48)$$

Proposition 4 (Dressed \uparrow -trees). *For $x \in E$, the process X^\uparrow is equal in law to X under $\mathbb{P}_{\delta_x}^\uparrow(\cdot | c_\emptyset(0) = \uparrow)$. Moreover, both are equal in law to a dressed $(\mathbb{P}^\uparrow, G^\uparrow)$ -MBP, say X^\uparrow , where the motion semigroup \mathbb{P}^\uparrow corresponds to the Markov process ξ on $E \cup \{\uparrow\}$ with probabilities*

$(\mathbf{P}_x^\uparrow, x \in E)$ given by (recalling that p is valued 0 on \dagger)

$$\frac{d\mathbf{P}_x^\uparrow}{d\mathbf{P}_x} \Big|_{\sigma(\xi_s, s \leq t)} = \frac{p(\xi_t)}{p(x)} \exp \left(- \int_0^t \frac{G[w](\xi_s)}{p(\xi_s)} ds \right), \quad t \geq 0, \quad (4.49)$$

and the branching generator is given by

$$G^\uparrow[f] = \frac{1}{p} (G[pf + w] - (1 - f)G[w]), \quad f \in L_\infty^{+,1}(E). \quad (4.50)$$

The dressing consists of additional particles, which are immigrated non-locally in space at the branch points of X^\uparrow , with each immigrated particle continuing to evolve as an independent copy of $(X^\downarrow, \mathbb{P}^\downarrow)$ from their respective space-point of immigration, such that the joint branching/immigration generator of type- \uparrow offspring and type- \downarrow immigrants is given by (4.48).

Proof of Proposition 4. We may think of $((x_i(t), c_i(t)), i \leq N_t), t \geq 0$, under \mathbb{P}^\dagger as a two-type branching process. To this end, let us write $N_t^\uparrow = \sum_{i=1}^{N_t} \mathbf{1}_{(c_i(t)=\uparrow)}$ and $N_t^\downarrow = N_t - N_t^\uparrow$, for $t \geq 0$. Define, for $f, g \in L_\infty^{+,1}(E)$,

$$u_t^\dagger[f, g](x) = \mathbb{E}_{\delta_x}^\dagger [\Pi_t[f, g] | c_\emptyset(0) = \dagger], \quad t \geq 0. \quad (4.51)$$

where, for $t \geq 0$,

$$\Pi_t[f, g] = \prod_{i=1}^{N_t^\uparrow} p(x_i^\uparrow(t)) f(x_i^\uparrow(t)) \prod_{j=1}^{N_t^\downarrow} w(x_j^\downarrow(t)) g(x_j^\downarrow(t)),$$

where

$$(x_i^\uparrow(t), i = 1, \dots, N_t^\uparrow) = (x_i(t) \text{ such that } c_i(t) = \uparrow, i \leq N_t)$$

and $(x_i^\downarrow(t), i = 1, \dots, N_t^\downarrow)$ is similarly defined.

We can break the expectation in the definition of $u_t^\dagger[f, g]$ over the first branching event, noting that until that moment, the initial ancestor is necessarily prolific. We have (again remembering $p(\dagger) = 0$)

$$\begin{aligned} u_t^\dagger[f, g](x) &= \frac{\mathbb{E}_{\delta_x} [\Pi_t[f, g] \mathbf{1}_{(c_\emptyset(0)=\dagger)}]}{\mathbb{P}_{\delta_x}(c_\emptyset(0) = \dagger)} \\ &= \frac{1}{p(x)} \mathbf{E}_x \left[p(\xi_t) f(\xi_t) e^{-\int_0^t \varsigma(\xi_u) du} \right] \\ &\quad + \frac{1}{p(x)} \mathbf{E}_x \left[\int_0^t p(\xi_s) \frac{\varsigma(\xi_s)}{p(\xi_s)} e^{-\int_0^s \varsigma(\xi_u) du} \right. \\ &\quad \left. \mathcal{E}_{\xi_s} \left[\sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \prod_{i \in I} p(x_i) u_{t-s}^\uparrow[f, g](x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i) u_{t-s}^\downarrow[g](x_i) \right] ds \right]. \end{aligned} \quad (4.52)$$

To help the reader interpret (4.52) better, we note that the first term on the right-hand side

comes from the event that no branching occurs up to time t , in which case the initial ancestor is positioned at ξ_t . Moreover, we have used the fact that $\mathbb{P}_{\delta_x}(c_\emptyset(0) = \uparrow | \mathcal{F}_t) = p(\xi_t)$. The second term is the consequence of a branching event occurring at time $s \in [0, t]$, at which point in time, the initial ancestor is positioned at ξ_s and thus has offspring scattered at $(x_i, i = 1 \cdots, N)$ according to \mathcal{P}_{ξ_s} . The contribution thereof from time s to t , can be either captured by $u_{t-s}^\uparrow[f, g]$, with probability p , if a given offspring is of type- \uparrow (thereby growing a tree of particles marked both \uparrow and \downarrow), or captured by $u_{t-s}^\downarrow[g]$, with probability w , if a given offspring is of type- \downarrow (thereby growing a tree of particles marked only with \downarrow). Hence projecting the expectation of $\Pi_t[f, g] \mathbf{1}_{(c_\emptyset = \uparrow)}$ onto the given configuration $(x_i, i = 1 \cdots, N)$ at time s , we get the sum inside the expectation with respect to \mathcal{P}_{ξ_s} , which caters for all the possible markings of the offspring of the initial ancestor, ensuring that at least one of them is \uparrow (which guarantees $c_\emptyset(0) = \uparrow$). In both expectations, the event of killing is accommodated for the fact that $p(\dagger) = f(\dagger) = \varsigma(\dagger) = 0$.

We may now substitute and (4.48) into (4.52) to get

$$\begin{aligned} & u_t^\uparrow[f, g](x) \\ &= \frac{1}{p(x)} \mathbf{E}_x \left[p(\xi_t) f(\xi_t) e^{-\int_0^t \varsigma(\xi_u) du} \right] \\ &+ \frac{1}{p(x)} \mathbf{E}_x \left[\int_0^t p(\xi_s) \frac{\varsigma(\xi_s)}{p(\xi_s)} e^{-\int_0^s \varsigma(\xi_u) du} \left[G^\uparrow[u_{t-s}^\uparrow[f, g], u_{t-s}^\downarrow[g]](\xi_s) + \varsigma^\uparrow(\xi_s) u_{t-s}^\uparrow[f, g](\xi_s) \right] ds \right]. \end{aligned} \quad (4.53)$$

Next, recalling the first equality in (4.44) that $\varsigma(x) = \varsigma^\uparrow(x) + G[w](x)/p(x)$, in each of the terms on the right-hand side of (4.53), we can exchange the exponential potential $\exp(-\int_0^t \varsigma(\xi_u) du)$ for the exponential potential $\exp(-\int_0^t G[w](\xi_u)/p(\xi_u) du)$ by transferring the the difference in the exponent to an additive potential (cf. Lemma 1.2, Chapter 4 in [10]). In this exchange, the term $\varsigma^\uparrow(\xi_s) u_{t-s}^\uparrow[f, g](\xi_s)$ is cancelled out on the right-hand side of (4.53). Then recalling the change of measure (4.49) that defines the semigroup \mathbf{P}^\uparrow , we get, on E ,

$$u_t^\uparrow[f, g] = \mathbf{P}_t^\uparrow[f] + \int_0^t \mathbf{P}_s^\uparrow \left[G^\uparrow[u_{t-s}^\uparrow[f, g], u_{t-s}^\downarrow[g]] \right] ds, \quad t \geq 0. \quad (4.54)$$

(Note, there is no need to define the object $\hat{\mathbf{P}}^\uparrow$ in the sprit of (4.26) as the semigroup \mathbf{P}^\uparrow is that of a conservative process on E .) This is the semigroup of a two-type MBP in which \downarrow -marked particles immigrate off an \uparrow -marked MBP. We have yet to verify however that the \uparrow -marked MBP is in fact the previously described $(\mathbf{P}^\uparrow, G^\uparrow)$ -MBP. In order to do this, we need to show that $G^\uparrow[f] = G^\uparrow[f, 1]$, for all $f \in L_\infty^{+,1}(E)$, where G^\uparrow was given in (4.50).

To this end, let us note two computational facts. First, for any $x \in E$,

$$1 = \mathcal{E}_x \left[\prod_{i=1}^N (p(x_i) + w(x_i)) \right] = \mathcal{E}_x \left[\sum_{I \subseteq \{1, \dots, N\}} \prod_{i \in I} p(x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i) \right], \quad (4.55)$$

so that

$$\varsigma^\uparrow(x) = \frac{\varsigma(x)}{p(x)} \mathcal{E}_x \left[\sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \prod_{i \in I} p(x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i) \right]. \quad (4.56)$$

Second, recalling (4.23), note

$$G[f](x) - G[g](x) = \varsigma(x) \left(\mathcal{E}_x \left[\prod_{j=1}^N f(x_j) \right] - \mathcal{E}_x \left[\prod_{j=1}^N g(x_j) \right] \right)$$

and $G[1](x) \equiv 0$. We thus have that, for $x \in E$,

$$\begin{aligned} G^\uparrow[f] &= G^\uparrow[f, 1](x) \\ &= \frac{\varsigma(x)}{p(x)} \mathcal{E}_x \left[\sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \prod_{i \in I} p(x_i) f(x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i) \right] \\ &\quad - f(x) \frac{\varsigma(x)}{p(x)} \mathcal{E}_x \left[\sum_{\substack{I \subseteq \{1, \dots, N\} \\ |I| \geq 1}} \prod_{i \in I} p(x_i) \prod_{i \in \{1, \dots, N\} \setminus I} w(x_i) \right] \\ &= \frac{\varsigma(x)}{p(x)} \mathcal{E}_x \left[\prod_{k=1}^N (p(x_k) f(x_k) + w(x_k)) - \prod_{k=1}^N w(x_k) \right] \\ &\quad - f(x) \frac{\varsigma(x)}{p(x)} \left[\prod_{k=1}^N (p(x_k) + w(x_k)) - \prod_{k=1}^N w(x_k) \right] \\ &= \frac{1}{p(x)} \{G[pf + w](x) - G[w](x)\} - f(x) \frac{1}{p(x)} \{G[1](x) - G[w](x)\} \\ &= \frac{1}{p(x)} \{G[pf + w](x) - (1 - f)G[w](x)\}, \end{aligned}$$

since $p(x) + w(x) = 1$, and this is just as required. \square

Theorem 14 (Skeletal decomposition). *We assume throughout that (M1) and (M2) are in force. Suppose that $\mu = \sum_{i=1}^n \delta_{x_i}$, for $n \in \mathbb{N}$ and $x_1, \dots, x_n \in E$. Then (X, \mathbb{P}_μ) is equal in law to*

$$\sum_{i=1}^n \left(\mathbf{B}_i X_t^{i, \uparrow} + (1 - \mathbf{B}_i) X_t^{i, \downarrow} \right), \quad t \geq 0, \quad (4.57)$$

where, for each $i = 1, \dots, n$, \mathbf{B}_i is an independent Bernoulli random variable with probability of success given by

$$p(x_i) := 1 - w(x_i) \quad (4.58)$$

and the processes $X^{i, \downarrow}$ and $X^{i, \uparrow}$ are independent copies of $(X, \mathbb{P}_{\delta_{x_i}}^\downarrow)$ and $(X, \mathbb{P}_{\delta_{x_i}}^\uparrow (\cdot | c_\emptyset(0) = \uparrow))$, respectively.

As alluded to previously, Theorem 14 pertains to a classical decomposition of branching trees in which the process (4.57) describes how the MBP divides into the genealogical lines of descent which are ‘prolific’ (surviving with probability p), in the sense that they create eternal subtrees which never leave the domain, and those which are ‘unsuccessful’ (dying with probability w), in the sense that they generate subtrees in which all genealogies die out.

Remark 8. It is an easy consequence of Theorem 14 that, for $t \geq 0$, the law of X_t^\uparrow conditional on $\mathcal{F}_t = \sigma(X_s, s \leq t)$, is equal to that of a Binomial point process with intensity $p(\cdot)X_t(\cdot)$. The latter, written $\text{BinPP}(pX_t)$, is an atomic random measure given by

$$\text{BinPP}(pX_t) = \sum_{i=1}^{N_t} B_i \delta_{x_i(t)},$$

where (we recall) that $X_t = \sum_{i=1}^{N_t} \delta_{x_i(t)}$, and B_i is a Bernoulli random variable with probability $p(x_i(t))$, $i = 1, \dots, N_t$.

Remark 9. It is also worth noting that the skeleton process X^\uparrow , given above, necessarily has at least one type- \uparrow offspring at each branch point, and indeed might have exactly one type- \uparrow offspring (although possibly with other simultaneous type- \downarrow immigrants). As such, an alternative way of looking at the type- \uparrow process would be to think of the skeleton of prolific individuals as a $(\mathbb{P}^\uparrow, G^\uparrow)$ -MBP with *at least two* type- \uparrow offspring at each branch point, with a modified motion \mathbb{P}^\uparrow in place of \mathbb{P}^\uparrow which integrates the event of a single type- \uparrow as an additional discontinuity in the movement. However, note these additional jumps are special in the sense as they are also potential points of simultaneous immigration of type- \downarrow particles, unlike other jumps corresponding to \mathbb{P}^\uparrow where there is no type- \downarrow immigration.

4.2.4 Combining \uparrow -trees and \downarrow -trees into the skeletal decomposition

Finally we are now ready to give the skeletal decomposition of (X, \mathbb{P}) .

Proof of Theorem 14. As previously, we may think of $((x_i(t), c_i(t)), i \leq N_t), t \geq 0$, as a two-type branching process under \mathbb{P}^\uparrow . A similar calculation to (4.33) gives us that, for $\nu = \sum_{i=1}^n \delta_{x_i}$ with $n \geq 1$ and $x_i \in E$, $i = 1, \dots, n$,

$$\mathbb{E}_\nu^\uparrow [\Pi_t[f, g]] = \sum_{I \subseteq \{1, \dots, n\}} \prod_{i \in I} p(x_i) u_t^\uparrow[f, g](x_i) \prod_{i \in \{1, \dots, n\} \setminus I} w(x_i) u_t^\downarrow[wg](x_i).$$

What this shows, together with the conditional version (4.33), is that the change of measure (4.33) (which is of course unity) is equivalent to a Doob h -transform on a two-type branching particle system (i.e. types $\{\uparrow, \downarrow\}$) where we consider the system after disregarding the marks. The effect of this Doob h -transform on type- \downarrow particles is that they generate $(\mathbb{P}^\downarrow, G^\downarrow)$ -MBPs, where as type- \uparrow particles generate a dressed $(\mathbb{P}^\uparrow, G^\uparrow)$ -MBP as described in Proposition 4. \square

4.2.5 Remarks on the skeletal decomposition for the NBP

The case of the skeletal decomposition for the NBP adds an additional layer of intricacy to the general picture given above. In this case, we have $E = D \times V$ with cemetery state \dagger that is entered when there is neutron capture (a neutron disappears in $D \times V$ without undergoing fission) or neutrons go to the physical boundary points $\{(r, v) : r \in \partial D \text{ and } v \cdot \mathbf{n}_r > 0\}$. It turns out that for the NBP, it is more convenient to view Theorem 14 in the spirit of Remark 9, i.e. we view the process X^\uparrow as a branching process that has at least two offspring at every branching event and whose movement corresponds to advection plus an extra discontinuity, which accounts for a branching event with one offspring.

To make this statement more precise, we first enforce the conditions of Theorem 12 in order to ensure (M1) and (M2) are satisfied. Indeed, on account of the inclusion $\{\zeta < \infty\} \subseteq \{W_\infty = 0\}$, we see that $w(x) \leq \mathbb{P}_{\delta_x}(W_\infty = 0)$, $r \in D, v \in V$. Recalling that W converges both almost surely as well as in $L^1(\mathbb{P})$ to its limit, we have that $\mathbb{P}_{\delta_x}(W_\infty = 0) < 1$ for $r \in D, v \in V$. This, combined with the fact that every particle may leave the bounded domain D directly without scattering or undergoing fission with positive probability, gives us that

$$e^{-\int_0^D \kappa_{r,v}^D \sigma(r+vs,v) ds} < w(r, v) < 1 \text{ for all } r \in D, v \in V. \quad (4.59)$$

Note that the lower bound is uniformly bounded away from 0 thanks to the maximal diameter of D , the minimal velocity v_{\min} (which, together uniformly upper bound $\kappa_{r,v}^D$) and the uniformly upper bounded rates of fission and scattering. The upper inequality becomes an equality for $r \in \partial D$ and $v \cdot \mathbf{n}_r > 0$.

Now, viewing the NBP X as a process with movement \mathbf{Q} and branching generator G , heuristically speaking, we can understand a little better the motions of X^\uparrow and X^\downarrow through the action of their generators. By considering only the leading order terms in small time (the process $(X_t, t \geq 0)$ is but a Markov chain), the action of the generator can be seen as the result of the limit

$$\overset{\leftarrow}{\mathbb{L}}f = \lim_{t \downarrow 0} \frac{1}{t} (\mathbf{Q}_t[f] - f), \quad (4.60)$$

for suitably smooth f (e.g. continuously differentiable within $L_\infty^+(D \times V)$). It is easy to show, and indeed known (cf. e.g. [6, 8]), that the action of the generator corresponding to \mathbf{Q} is given by

$$\overset{\leftarrow}{\mathbb{L}}f(r, v) = v \cdot \nabla f(r, v) + \int_V (f(r, v') - f(r, v)) \sigma_{\mathbf{s}}(r, v) \pi_{\mathbf{s}}(r, v, v') dv', \quad (4.61)$$

for $f \in L_\infty^+(D \times V)$ such that ∇f is well defined (here ∇ is assumed to act on the spatial variable r). We emphasise again that, in view of Remark 9, this corresponds to motion plus a branching event with one offspring (or scattering).

The change of measure (4.35) induces a generator action given by

$$\begin{aligned}
\check{\mathbb{L}}^\downarrow f(r, v) &= \frac{1}{w(r, v)} \check{\mathbb{L}}(wf)(r, v) + f(r, v) \frac{G[w]}{w}(r, v) \\
&= v \cdot \nabla f(r, v) + \int_V (f(r, v') - f(r, v)) \sigma_{\mathbf{s}}(r, v) \frac{w(r, v')}{w(r, v)} \pi_{\mathbf{s}}(r, v, v') dv' \\
&\quad + f(r, v) \left(\frac{\check{\mathbb{L}}w}{w} + \frac{G[w]}{w} \right) (r, v) \\
&= v \cdot \nabla f(r, v) + \int_V (f(r, v') - f(r, v)) \sigma_{\mathbf{s}}(r, v) \frac{w(r, v')}{w(r, v)} \pi_{\mathbf{s}}(r, v, v') dv', \quad (4.62)
\end{aligned}$$

where the fact that the right-hand side of (4.35) is a martingale will lead to $\check{\mathbb{L}}w + G[w] = 0$.

In other words, our heuristic reasoning above shows that the motion on the \downarrow -marked tree is tantamount to a w -tilting of the scattering kernel. This tilting favours scattering in a direction where extinction becomes more likely, and as such, $\check{\mathbb{L}}^\downarrow$ encourages \downarrow -marked trees to become extinct ‘quickly’.

Almost identical reasoning shows that the change of measure (4.49) has generator with action

$$\begin{aligned}
\check{\mathbb{L}}^\uparrow f(r, v) &= \frac{1}{p(r, v)} \check{\mathbb{L}}(pf)(r, v) - f(r, v) \frac{G[w]}{p}(r, v) \\
&= v \cdot \nabla f(r, v) + \int_V (f(r, v') - f(r, v)) \sigma_{\mathbf{s}}(r, v) \frac{p(r, v')}{p(r, v)} \pi_{\mathbf{s}}(r, v, v') dv', \quad (4.63)
\end{aligned}$$

for suitably smooth f , where we have again used $\check{\mathbb{L}}w + G[w] = 0$ and left the calculations that the second equality from the first as an exercise for the reader. One sees again a p -tilting of the scattering kernel, and hence L^\uparrow rewards scattering in directions that ‘enable survival’. Note, moreover for regions of $D \times V$ for which $p(r, v)$ can be come arbitrarily small (corresponding to a small probability of survival), the scattering rate also becomes very large, and hence L^\uparrow ‘urgently’ scatters particles away from such regions.

4.2.6 Remarks on BBM

On account of the fact that we have stated Theorem 14 for a relatively general MBP with non-local branching, it is worth pointing to the known example of a BBM in a strip that has previously been worked out in detail in [14]. This model has the features that \mathbf{P} is that of a Brownian motion with drift μ killed on existing an interval $[0, K]$, so that $\check{\mathbb{L}} = (1/2)d^2/dx^2 + \mu d/dx$, the branching rate ς is constant (not spatially dependent) and the offspring distribution is concentrated at the point of death of each particle. As such, the generator G in (4.23) takes the simpler form

$$G[\theta] = \varsigma \mathcal{E} (\theta^N - \theta) \quad (4.64)$$

where it suffices to take θ as a number in $(0, 1)$, rather than a function, as there is no spatial dependency. The extinction probability now solves the differential equation

$$\frac{1}{2} \frac{dw}{dx^2} + \mu \frac{dw}{dx} + G[w] = 0 \text{ on } (0, K) \text{ with } w(0) = w(K) = 1.$$

In order for survival to occur with positive probability, it is required that the leading eigenvalue of the mean semigroup associated to the branching process, which is $\lambda_* := (m-1)\zeta - \mu^2/2 - \pi^2/2K$, must satisfy $\lambda_* > 0$, where $m = \sum_{k=0}^{\infty} kp_k$ is the mean number of offspring. Note, the mean semigroup is the analogue of (4.10) and the leading eigenvalue plays precisely the role of λ_* in Theorem 11 for the NTE.

For the \downarrow process, writing G^\downarrow in (4.36) in a similar format to (4.64), it is straightforward to verify that it agrees with the branching generator stipulated in analysis of the red tree given in [14].

However, for the \uparrow process, this model also takes the point of view described in Remark 9. Indeed, it is straightforward to show that the branching generator for the blue tree in [14] agrees with G^\uparrow given in Remark 9 and the ‘discontinuity’ associated with a birth of one offspring is appended to the motion. However, since this model only has local branching and the movement is a Brownian motion, this does not actually change the motion. On the other hand, this choice does affect the overall process X^\uparrow and leads to two types of immigration of red trees onto the blue tree: immigration at branch points and immigration along the trajectory, with the latter immigration occurring at the points corresponding to a ‘birth of one offspring’.

When, additionally, the interval $[0, K]$ is replaced by \mathbb{R} , the extinction probability w is no longer spatially dependent and is a simple solution of $G[w] = 0$. Assuming that $w \in (0, 1)$, it is easy to see that $\overleftarrow{L}^\uparrow$ and $\overleftarrow{L}^\downarrow$ are both equal to \overleftarrow{L} and the skeletal decomposition is nothing more than the original skeletal decomposition for Galton–Watson processes (albeit in continuous time) given in the book of Harris [16].

4.3 SLLN on the skeleton

Our aim is to use the skeletal decomposition of the neutron branching process to prove Theorem 13 by first stating and proving the analogous result for X^\uparrow . Hence, in what follows, we will assume (H1), (H2)*, (H3)* and (H4) hold. Before continuing to the proof, let us consider a useful identity. For a suitable $g \in L_\infty(D \times V)$ and $t \geq 0$, we have from Theorem 14 (cf. Remark 8) that

$$\mathbb{E}_{\delta_{(r,v)}^\uparrow} [\langle g, X_t^\uparrow \rangle] = \mathbb{E}_{\delta_{(r,v)}^\downarrow} [\langle g, pX_t \rangle | c_\emptyset(0) = \uparrow] = \frac{1}{p(r,v)} \mathbb{E}_{\delta_{(r,v)}} [\langle gp, X_t \rangle]. \quad (4.65)$$

We can use this identity to show that λ_* is also an eigenvalue for the linear semigroup of X^\uparrow , as well as to compute the associated left and right eigenfunctions (in a similar sense to (4.18)). Our first claim is that the right eigenfunction is given by φ/p . Indeed, for $(r, v) \in D \times V$,

due to the above computation,

$$\mathbb{E}_{\delta_{(r,v)}^\uparrow}[\langle \varphi/p, X_t^\uparrow \rangle] = \frac{\mathbb{E}_{\delta_{(r,v)}}[\langle \varphi, X_t \rangle]}{p(r,v)} = e^{\lambda_* t} \frac{\varphi(r,v)}{p(r,v)}. \quad (4.66)$$

For the left eigenfunction, again using (4.65), we have

$$\langle \tilde{\varphi} p, \mathbb{E}_{\delta_{(r,v)}^\uparrow}[\langle g, X_t^\uparrow \rangle] \rangle = \langle \tilde{\varphi} p, \mathbb{E}_{\delta_{(r,v)}}[\langle gp, X_t \rangle] / p(\cdot) \rangle = \langle \tilde{\varphi}, \mathbb{E}_{\delta_{(r,v)}}[\langle gp, X_t \rangle] \rangle = e^{\lambda_* t} \langle \tilde{\varphi} p, g \rangle. \quad (4.67)$$

Hence $\tilde{\varphi} p$ is the corresponding left eigenfunction with eigenvalue $e^{\lambda_* t}$.

It now follows by similar arguments to those given in [17] that

$$W_t^\uparrow := e^{-\lambda_* t} \frac{\langle \varphi/p, X_t^\uparrow \rangle}{\langle \varphi/p, \mu \rangle}, \quad t \geq 0, \quad (4.68)$$

is a positive martingale under \mathbb{P}_μ^\uparrow for $\mu \in \mathcal{M}(D \times V)$, and hence has a finite limit, which we denote W_∞^\uparrow .

A second useful fact that we will use is the following result.

Lemma 10. *There exists a constant $C \in (0, \infty)$ such that $\sup_{r \in D, v \in V} \varphi(r, v) / p(r, v) < C$.*

Proof. Let us introduce the family of measures $\mathbb{P}^\varphi := (\mathbb{P}_\mu^\varphi, \mu \in \mathcal{M}(D \times V))$, where

$$\left. \frac{d\mathbb{P}_\mu^\varphi}{d\mathbb{P}_\mu} \right|_{\mathcal{F}_t} = W_t, \quad t \geq 0, \quad (4.69)$$

We start by noting that, for all $r \in D, v \in V$, $p(r, v) = 1 - \mathbb{P}_{\delta_{(r,v)}}(\zeta < \infty) = \mathbb{P}_{\delta_{(r,v)}}(X \text{ survives})$, where ζ is the lifetime of X defined in (4.27). Taking account of (4.69), we can thus write, with the help of Fatou's Lemma and Jensen's inequality,

$$\begin{aligned} p(r, v) &= \lim_{t \rightarrow \infty} \mathbb{P}_{\delta_{(r,v)}}(t < \zeta) \\ &= \lim_{t \rightarrow \infty} \mathbb{E}_{\delta_{(r,v)}^\varphi} \left[\frac{1}{W_t} \mathbf{1}_{(t < \zeta)} \right] \\ &\geq \mathbb{E}_{\delta_{(r,v)}^\varphi} [1/W_\infty] \\ &\geq 1/\mathbb{E}_{\delta_{(r,v)}^\varphi} [W_\infty], \end{aligned}$$

where we note that the indicator is dropped in the first inequality as, from Lemma 6.1 in [17], the process (X, \mathbb{P}^φ) is immortal.

From equations (10.1) and (10.3) in [17], it has already been shown that

$$\mathbb{E}_{\delta_{(r,v)}^\varphi} [W_\infty] = \lim_{t \rightarrow \infty} \mathbb{E}_{\delta_{(r,v)}} [W_t^2] \leq c \int_0^\infty e^{-2\lambda_* t} \frac{\psi_t[1](r, v)}{\varphi(r, v)} dt, \quad (4.70)$$

for some constant $c \in (0, \infty)$. Taking account of Theorem 11, which tells us that

$$\lim_{t \rightarrow \infty} e^{-\lambda_* t} \psi_t[1](r, v) = \langle 1, \tilde{\varphi} \rangle \varphi(r, v) \leq \|\varphi\|_\infty \langle 1, \tilde{\varphi} \rangle,$$

we deduce that there exists a constant $C \in (0, \infty)$, which does not depend on $(r, v) \in D \times V$, such that

$$p(r, v) \geq \frac{\varphi(r, v)}{C}.$$

The result now follows. \square

We are now in a position to state and prove a strong law for the skeleton X^\uparrow .

Theorem 15. *For all non-negative and directionally continuous g (in the sense that $\lim_{s \downarrow 0} g(r + vs, v) = g(r, v)$ for all $r \in D, v \in V$) such that, for some constant $c > 0$, $g \leq c\varphi/p$,*

$$\lim_{t \rightarrow \infty} e^{-\lambda_* t} \langle g, X_t^\uparrow \rangle = \langle g, \tilde{\varphi} p \rangle \langle \varphi/p, \mu \rangle W_\infty^\uparrow. \quad (4.71)$$

\mathbb{P}_μ^\uparrow -almost surely for $\mu \in \mathcal{M}(D \times V)$

We prove this theorem by breaking it up into several parts. Starting with the following lemma, we first prove that Theorem 15 holds along lattice times. Our proofs are principally by techniques that have been used a number of times in the literature, developed by [1, 13, 4, 3] amongst others. Just before we state the next lemma, it will be convenient to quickly introduce the notation $\mathcal{F}_t^\uparrow = \sigma(X_s^\uparrow : s \leq t)$, $t \geq 0$.

Lemma 11. *Fix $\delta > 0$. For non-negative bounded functions $g \in L_\infty^+(D \times V)$, define*

$$U_t = e^{-\lambda_* t} \langle g\varphi/p, X_t^\uparrow \rangle, \quad t \geq 0. \quad (4.72)$$

Then, for any non-decreasing sequence $(m_n)_{n \geq 0}$ with $m_0 > 0$ and $(r, v) \in D \times V$,

$$\lim_{n \rightarrow \infty} |U_{(m_n+n)\delta} - \mathbb{E}^\uparrow[U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow]| = 0, \quad \mathbb{P}_{\delta(r,v)}^\uparrow \text{-a.s.} \quad (4.73)$$

Proof. By the Borel-Cantelli lemma, it is sufficient to prove that for each $(r, v) \in D \times V$ and all $\varepsilon > 0$,

$$\sum_{n \geq 1} \mathbb{P}_{\delta(r,v)}^\uparrow \left(|U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow]| > \varepsilon \right) < \infty.$$

To this end, note that Markov's inequality gives

$$\mathbb{P}_{\delta(r,v)}^\uparrow \left(|U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow]| > \varepsilon \right) \leq \varepsilon^{-2} \mathbb{E}_{\delta(r,v)}^\uparrow \left(|U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow]|^2 \right).$$

Hence, let us consider the term in the conditional expectation on the right-hand side above. First note that

$$U_{(m_n+n)\delta} - \mathbb{E}^\uparrow[U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow] = \sum_{i=1}^{N_{n\delta}} e^{-n\delta\lambda_*} (U_{m_n\delta}^{(i)} - \mathbb{E}^\uparrow[U_{m_n\delta}^{(i)} | \mathcal{F}_{n\delta}^\uparrow]), \quad (4.74)$$

where, conditional on $\mathcal{F}_{n\delta}^\uparrow$, $Z_i = U_{m_n\delta}^{(i)} - \mathbb{E}^\uparrow(U_{m_n\delta}^{(i)} | \mathcal{F}_{n\delta}^\uparrow)$ are independent with $\mathbb{E}[Z_i] = 0$. The formula for the variance of sums of zero mean independent random variables together with the inequality $|a + b|^2 \leq 2(|a|^2 + |b|^2)$, we get

$$\begin{aligned} & \mathbb{E}^\uparrow(|U_{(m_n+n)\delta} - \mathbb{E}[U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow]|^2 | \mathcal{F}_{n\delta}^\uparrow) \\ &= \sum_{i=1}^{N_{n\delta}} e^{-2\lambda_* n\delta} \mathbb{E}^\uparrow \left[\left| U_{m_n\delta}^{(i)} - \mathbb{E}^\uparrow[U_{m_n\delta}^{(i)} | \mathcal{F}_{n\delta}^\uparrow] \right|^2 \middle| \mathcal{F}_{n\delta}^\uparrow \right] \\ &\leq \sum_{i=1}^{N_{n\delta}} e^{-2\lambda_* n\delta} \mathbb{E}^\uparrow \left[4(|U_{m_n\delta}^{(i)}|^2 + |\mathbb{E}^\uparrow[U_{m_n\delta}^{(i)} | \mathcal{F}_{n\delta}^\uparrow]|^2) | \mathcal{F}_{n\delta}^\uparrow \right] \\ &\leq 4 \sum_{i=1}^{N_{n\delta}} e^{-2\lambda_* n\delta} \mathbb{E}^\uparrow \left[|U_{m_n\delta}^{(i)}|^2 | \mathcal{F}_{n\delta}^\uparrow \right], \end{aligned}$$

where we have used Jensen's inequality again in the final inequality. Hence, with $\{(R_i(n\delta), \Upsilon_i(n\delta)) : i = 1, \dots, N_{n\delta}\}$ describing the configurations of the particles at time $N_{n\delta}$ in X^\uparrow , we have

$$\begin{aligned} & \sum_{n=1}^{\infty} \mathbb{E}^\uparrow \left[|U_{(m_n+n)\delta} - \mathbb{E}^\uparrow(U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow)|^2 \right] \\ &\leq 4 \sum_{n=1}^{\infty} e^{-2\lambda_* n\delta} \mathbb{E}_{\delta(r,v)}^\uparrow \left[\sum_{i=1}^{N_{n\delta}} \mathbb{E}_{\delta(R_i(n\delta), \Upsilon_i(n\delta))}^\uparrow [U_{m_n\delta}^2] \right] \\ &\leq 4 \|g\|_\infty^2 \sum_{n=1}^{\infty} e^{-2\lambda_* n\delta} \mathbb{E}_{\delta(r,v)}^\uparrow \left[\sum_{i=1}^{N_{n\delta}} \frac{\varphi(R_i(n\delta), \Upsilon_i(n\delta))^2}{p(R_i(n\delta), \Upsilon_i(n\delta))^2} \mathbb{E}_{\delta(R_i(n\delta), \Upsilon_i(n\delta))}^\uparrow [(W_{m_n\delta}^\uparrow)^2] \right], \quad (4.75) \end{aligned}$$

where the final inequality was obtained by noting that, from the definitions of U_t and W_t^\uparrow , we have

$$\mathbb{E}_{\delta(r,v)}^\uparrow [U_t^2] \leq \|g\|_\infty^2 \frac{\varphi(r,v)^2}{p(r,v)^2} \mathbb{E}_{\delta(r,v)}^\uparrow [(W_t^\uparrow)^2].$$

Due to Theorem 14, in particular Remark 8, and the calculation leading to (4.65), we have, for all $t \geq 0$,

$$\begin{aligned} \mathbb{E}_{\delta(r,v)}^\uparrow [(W_t^\uparrow)^2] &= \frac{e^{-2\lambda_* t}}{(\varphi(r,v)/p(r,v))^2} \mathbb{E}_{\delta(r,v)}^\uparrow [\langle \varphi/p, X_t^\uparrow \rangle^2] \\ &= \frac{e^{-2\lambda_* t}}{(\varphi(r,v)/p(r,v))^2} \mathbb{E}_{\delta(r,v)}^\uparrow [\langle \varphi/p, \text{BinPP}(pX_t) \rangle^2 | c_\emptyset(0) = \uparrow] \\ &\leq \frac{p(r,v)^2}{\varphi(r,v)^2} \left\{ e^{-2\lambda_* t} \mathbb{E}_{\delta(r,v)} [\langle \varphi^2/p, X_t \rangle] / p(r,v) + e^{-2\lambda_* t} \mathbb{E}_{\delta(r,v)} [\langle \varphi, X_t \rangle^2] / p(r,v) \right\} \\ &\leq C \left(\frac{e^{-\lambda_* t}}{\varphi(r,v)} + \mathbb{E}_{\delta(r,v)} [W_t^2] \right) p(r,v) \\ &\leq C \frac{p(r,v)}{\varphi(r,v)} \left(\varphi(r,v) \mathbb{E}_{\delta(r,v)} [W_t^2] + 1 \right) \end{aligned} \quad (4.76)$$

where we have used Lemma 10 in the second inequality. From Corollary 5.3 of [17], more

precisely from its proof, we know that $\mathbb{E}_{\delta(r,v)}[\sup_{t \geq 0} W_t^2] < \infty$. Hence we have from Doob's maximal inequality that, for each fixed $t \geq 0$,

$$\begin{aligned} \mathbb{E}_{\delta(r,v)}^\uparrow \left[(W_t^\uparrow)^2 \right] &\leq \mathbb{E}_{\delta(r,v)}^\uparrow \left[\sup_{s \geq 0} (W_s^\uparrow)^2 \right] \\ &\leq \limsup_{s \rightarrow \infty} 4\mathbb{E}_{\delta(r,v)}^\uparrow \left[(W_s^\uparrow)^2 \right] \\ &\leq 4C \frac{p(r,v)}{\varphi(r,v)} \left(\varphi(r,v) \mathbb{E}_{\delta(r,v)} [W_\infty^2] + 1 \right) \\ &\leq 4C \frac{p(r,v)}{\varphi(r,v)} (C' + 1) < \infty \end{aligned} \quad (4.77)$$

for some constant C' which does not depend on (r,v) , where we have used (4.70). (Note (4.77) implies that W^\uparrow is an $L_2(\mathbb{P}^\uparrow)$ -convergent martingale.)

Substituting the estimate (4.76) back into (4.75) and making use of the uniform boundedness of φ , we get

$$\begin{aligned} &\sum_{n=1}^{\infty} \mathbb{E}^\uparrow \left[|U_{(m_n+n)\delta} - \mathbb{E}^\uparrow(U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow)|^2 \right] \\ &\leq K \|g\|_\infty^2 \sum_{n=1}^{\infty} e^{-2\lambda_* n\delta} \mathbb{E}_{\delta(r,v)}^\uparrow \left[\sum_{i=1}^{N_{n\delta}} \frac{\varphi(R_i(n\delta), \Upsilon_i(n\delta))}{p(R_i(n\delta), \Upsilon_i(n\delta))} \right], \end{aligned} \quad (4.78)$$

for some constant $K \in (0, \infty)$. Now the fact that φ/p is an eigenfunction for the linear semigroup of X^\uparrow , we get

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{E}^\uparrow \left[|U_{(m_n+n)\delta} - \mathbb{E}^\uparrow(U_{(m_n+n)\delta} | \mathcal{F}_{n\delta}^\uparrow)|^2 \right] &\leq K \|g\|_\infty^2 \sum_{n \geq 1} e^{-2\lambda_* n\delta} \mathbb{E}_{\delta(r,v)}^\uparrow [\langle \varphi/p, X_{n\delta}^\uparrow \rangle] \\ &= K \|g\|_\infty^2 \frac{\varphi(r,v)}{p(r,v)} \sum_{n \geq 1} e^{-\lambda_* n\delta} < \infty. \end{aligned} \quad (4.79)$$

Taking expectations one more time with respect to $\mathbb{P}_{\delta(r,v)}^\uparrow$ and appealing to Fubini's Theorem to exchange the sum and expectation on the left hand side of (4.79) completes the proof of the lemma. \square

It is worth noting that a small corollary falls out of the above proof, which will be useful later on.

Corollary 3. *We have $\sup_{t \geq 0} W_t^\uparrow$ is square integrable and hence W^\uparrow converges in $L_2(\mathbb{P}^\uparrow)$.*

Proof of Theorem 15 (lattice sequences). We have already noted that

$$\mathbb{E}_{\delta(r,v)}^\uparrow \left[U_{t+s} | \mathcal{F}_t^\uparrow \right] = \sum_{i=1}^{N_t} e^{-\lambda_* t} U_s^{(i)},$$

where, given \mathcal{F}_t^\uparrow , the $U_s^{(i)}$ are independent and equal in distribution to U_s under $\mathbb{P}_{\delta_{(R_i(t), \Upsilon_i(t))}}^\uparrow$ and $\{(R_i(t), \Upsilon_i(t)) : i = 1, \dots, N_t\}$ describes the configuration of X^\uparrow at time $t \geq 0$. Hence, once again using (4.65), the many-to-one formula and the spine change of measure, we have

$$\begin{aligned}
\mathbb{E}_{\delta_{(r,v)}^\uparrow} \left[U_{t+s} | \mathcal{F}_t^\uparrow \right] &= \sum_{i=1}^{N_t} e^{-\lambda_* t} \mathbb{E}_{\delta_{(R_i(t), \Upsilon_i(t))}^\uparrow} \left[e^{-\lambda_* s} \langle g\varphi/p, X_s^\uparrow \rangle \right] \\
&= \sum_{i=1}^{N_t} e^{-\lambda_* t} \frac{\mathbb{E}_{\delta_{(R_i(t), \Upsilon_i(t))}^\uparrow} [e^{-\lambda_* s} \langle g\varphi, X_s \rangle]}{p(R_i(t), \Upsilon_i(t))} \\
&= \sum_{i=1}^{N_t} e^{-\lambda_* t} \frac{e^{-\lambda_* s} \psi_s[\varphi g](R_i(t), \Upsilon_i(t))}{p(R_i(t), \Upsilon_i(t))} \\
&= \frac{p(r, v)}{\varphi(r, v)} \langle g\varphi, \tilde{\varphi} \rangle W_t^\uparrow \\
&\quad + \sum_{i=1}^{N_t} e^{-\lambda_* t} \left(e^{-\lambda_* s} \frac{\psi_s[\varphi g](R_i(t), \Upsilon_i(t))}{\varphi(R_i(t), \Upsilon_i(t))} - \langle g\varphi, \tilde{\varphi} \rangle \right) \frac{\varphi(R_i(t), \Upsilon_i(t))}{p(R_i(t), \Upsilon_i(t))}.
\end{aligned} \tag{4.80}$$

Appealing to Theorem 11, we can pick s sufficiently large so that, for any given $\varepsilon > 0$,

$$\|e^{-\lambda_* s} \varphi^{-1} \psi_s[\varphi g] - \langle \tilde{\varphi}, \varphi g \rangle\|_\infty < \varepsilon. \tag{4.81}$$

Combining this with (4.80) yields

$$\lim_{t \rightarrow \infty} \left| \mathbb{E}_{\delta_{(r,v)}^\uparrow} [U_{t+s} | \mathcal{F}_t^\uparrow] - W_\infty^\uparrow \langle \varphi g, \tilde{\varphi} \rangle \frac{p(r, v)}{\varphi(r, v)} \right| = 0. \tag{4.82}$$

The above combined with the conclusion of Lemma 11 gives the conclusion of Theorem 13 along lattice sequences. \square

We now make the transition from lattice times to continuous times.

Proof of Theorem 15 (full sequence). For $\varepsilon > 0$ and $(r, v) \in D \times V$, define

$$\Omega_\varepsilon(r, v) := \left\{ (r', v') \in D \times V : g(r', v') \frac{\varphi(r', v')}{p(r', v')} \geq (1 + \varepsilon)^{-1} g(r, v) \frac{\varphi(r, v)}{p(r, v)} \right\}.$$

If we consider the equation (4.30) for the special setting of the NBP, we can decompose it over the first scatter event, rather than the first fission event, from which we will obtain

$$w(r, v) = \hat{U}_t[w](r, v) + \int_0^t \mathbb{U}_s \left[\overleftarrow{\mathbb{S}} w + G[w] \right] (r, v) ds, \quad t \geq 0, x \in E,$$

where the semigroup $(\mathbb{U}_t, t \geq 0)$ was defined in (4.4), $(\hat{U}_t, t \geq 0)$ was defined in (4.14), and the scattering operator $\overleftarrow{\mathbb{S}}$ was defined in (4.5). This implies that, for a given $r \in D$ and $v \in V$, $w(r + vt, v)$, and hence $p(r + vt, v)$, are continuous for all t sufficiently small. Similarly noting

that $\psi_t[\varphi] = e^{\lambda_* t} \varphi$, from (4.3), we can also deduce a similar continuity property of φ . Hence, together with the assumed directional continuity of g , for each $r \in D$, $v \in V$ and $\varepsilon \ll 1$, there exists a δ_ε such that $(r + vt, v) \in \Omega_\varepsilon(r, v)$ for all $t \leq \delta_\varepsilon$.

Next, for each $\delta > 0$ define

$$\Xi^{\delta, \varepsilon}(r, v) := \mathbf{1}_{\{\text{supp}(X_t^\uparrow) \subset \Omega_\varepsilon(r, v) \text{ for all } t \in [0, \delta]\}}, \quad (r, v) \in D \times V,$$

and let $\eta^{\delta, \varepsilon}(r, v) = \mathbb{E}_{\delta(r, v)}^\uparrow[\Xi^{\delta, \varepsilon}(r, v)] \leq 1$. Appealing to Fatou's Lemma and the continuity properties discussed above, we have, for $\varepsilon \ll 1$,

$$\begin{aligned} \liminf_{\delta \downarrow 0} \eta^{\delta, \varepsilon}(r, v) &\geq \mathbb{E}_{\delta(r, v)}^\uparrow \left[\liminf_{\delta \downarrow 0} \mathbf{1}_{\{\text{supp}(X_t^\uparrow) \subset \Omega_\varepsilon(r, v) \text{ for all } t \in [0, \delta]\}} \right] \\ &= \mathbb{E}_{\delta(r, v)}^\uparrow \left[\lim_{\delta \downarrow 0} \mathbf{1}_{\{(r+vt, v) \in \Omega_\varepsilon(r, v) \text{ for all } t \in [0, \delta]\}} \right] \\ &= 1. \end{aligned}$$

Since we can effectively see the skeleton as producing at least two offspring at every fission event (see Remark 9)⁸, it follows that if $t \in [n\delta, (n+1)\delta)$ then,

$$\begin{aligned} &e^{-\lambda_* t} \langle g\varphi/p, X_t^\uparrow \rangle \\ &\geq \frac{e^{-\delta}}{(1+\varepsilon)} \sum_{i=1}^{N_{n\delta}} e^{-\lambda_* n\delta} g(R_i(n\delta), \Upsilon_i(n\delta)) \frac{\varphi(R_i(n\delta), \Upsilon_i(n\delta))}{p(R_i(n\delta), \Upsilon_i(n\delta))} \Xi^{\delta, \varepsilon}(R_i(n\delta), \Upsilon_i(n\delta)). \end{aligned} \quad (4.83)$$

If we denote the summation on the right-hand side of the above equation by $\tilde{U}_{n\delta}(r, v)$, and assume that $\text{supp}(g)$ is compactly embedded in D , then we can apply similar arguments to those given in the proof of Lemma 11 together with (4.65) to show that

$$\begin{aligned} &\sum_{n=1}^{\infty} \mathbb{E}_{\delta(r, v)}^\uparrow \left[|\tilde{U}_{n\delta} - \mathbb{E}^\uparrow[\tilde{U}_{n\delta} | \mathcal{F}_{n\delta}^\uparrow]|^2 \right] \\ &\leq C \sum_{n=1}^{\infty} e^{-\lambda_* n\delta q} \mathbb{E}_{\delta(r, v)}^\uparrow \left[\sum_{i=1}^{N_{n\delta}} g(R_i(n\delta), \Upsilon_i(n\delta))^2 \frac{\varphi(R_i(n\delta), \Upsilon_i(n\delta))^2}{p(R_i(n\delta), \Upsilon_i(n\delta))^2} \eta^{\delta, \varepsilon}(R_i(n\delta), \Upsilon_i(n\delta)) \right] \\ &\leq C \sum_{n=1}^{\infty} e^{-2\lambda_* n\delta} \mathbb{E}_{\delta(r, v)}^\uparrow \left[\langle (g\varphi/p)^2, X_{n\delta}^\uparrow \rangle \right] \\ &\leq \frac{C}{p(r, v)} \sum_{n=1}^{\infty} e^{-2\lambda_* n\delta} \mathbb{E}_{\delta(r, v)} \left[\langle (g\varphi)^2 p^{-1}, X_{n\delta} \rangle \right] \\ &= \frac{C}{p(r, v)} \sum_{n=1}^{\infty} e^{-2\lambda_* n\delta} \psi_{n\delta}[(g\varphi)^2 p^{-1}](r, v). \end{aligned} \quad (4.84)$$

Note in particular that the compact embedding of the support of g in $D \times V$ together with Lemma 10, the fact that $p \leq 1$, φ belongs to $L_\infty^+(D \times V)$ and is bounded away from 0 on

⁸Although a subtle point in the argument, this is fundamentally the reason why the skeletal decomposition is needed and makes the proof much easier than otherwise.

compactly embedded subsets of $D \times V$ ensures that $(g\varphi)^2 p^{-1}$ is uniformly bounded away from 0 and ∞ and hence, taking account of the conclusion of Theorem 11, the expectation on the right-hand side of (4.84) is finite.

Noting that

$$\mathbb{E}^\uparrow[\tilde{U}_{n\delta} | \mathcal{F}_{n\delta}^\uparrow] = e^{-\lambda_* n\delta} \langle g\varphi\eta^{\delta,\varepsilon}/p, X_{n\delta}^\uparrow \rangle,$$

the consequence of (4.84), when taken in the light of the Borel-Cantelli Lemma and the already proved limit (4.71) on lattice times, means that, $\mathbb{P}_{\delta(r,v)}$ -almost surely,

$$\liminf_{t \rightarrow \infty} e^{-\lambda_* t} \langle g\varphi/p, X_t^\uparrow \rangle \geq \frac{e^{-\delta}}{1+\varepsilon} \langle g\varphi\eta^{\delta,\varepsilon}/p, \tilde{\varphi}p \rangle W_\infty^\uparrow \frac{\varphi(r,v)}{p(r,v)}.$$

Letting $\delta \downarrow 0$ with the help of Fatou's Lemma and then $\varepsilon \downarrow 0$ in the above inequality yields

$$\liminf_{t \rightarrow \infty} e^{-\lambda_* t} \langle g\varphi/p, X_t^\uparrow \rangle \geq \langle g\varphi, \tilde{\varphi} \rangle W_\infty^\uparrow \frac{\varphi(r,v)}{p(r,v)}, \quad (4.85)$$

$\mathbb{P}_{\delta(r,v)}$ -almost surely. Now replacing g by hp/φ , ensuring still that the support of h is compactly embedded in $D \times V$, so that hp/φ is uniformly bounded away from 0 and ∞ , the lower bound (4.85) yields

$$\liminf_{t \rightarrow \infty} e^{-\lambda_* t} \langle h, X_t^\uparrow \rangle \geq \langle h, \tilde{\varphi}p \rangle W_\infty^\uparrow \frac{\varphi(r,v)}{p(r,v)}. \quad (4.86)$$

We can push (4.86) a little bit further by removing the requirement that the support of h is compactly embedded in $D \times V$. Indeed, suppose that, for $n \geq 1$, $h_n = h\mathbf{1}_{B_n}$, where $h \leq c\varphi/p$ for some constant $c > 0$ and B_n is an increasing sequence of compactly embedded domains in $D \times V$, such that $\cup_{n \geq 1} B_n = D \times V$. Then (4.86) and together with monotonicity gives us

$$\begin{aligned} \liminf_{t \rightarrow \infty} e^{-\lambda_* t} \langle h, X_t^\uparrow \rangle &\geq \lim_{n \rightarrow \infty} \liminf_{t \rightarrow \infty} e^{-\lambda_* t} \langle h_n, X_t^\uparrow \rangle \\ &\geq \lim_{n \rightarrow \infty} \langle h_n, \tilde{\varphi}p \rangle W_\infty^\uparrow \frac{\varphi(r,v)}{p(r,v)} \\ &= \langle h, \tilde{\varphi}p \rangle W_\infty^\uparrow \frac{\varphi(r,v)}{p(r,v)} \end{aligned} \quad (4.87)$$

$\mathbb{P}_{\delta(r,v)}$ -almost surely.

To complete the proof of Theorem 15 it now suffices to show that, $\mathbb{P}_{\delta(r,v)}$ -almost surely, $\limsup_{t \rightarrow \infty} e^{-\lambda_* t} \langle g, X_t^\uparrow \rangle \leq \langle g, \tilde{\varphi}p \rangle W_\infty^\uparrow \varphi(r,v)/p(r,v)$. To this end note that, for $0 \leq g \leq c\varphi/p$,

for some constant $c > 0$ (which, without loss of generality, we may take equal to 1),

$$\begin{aligned}
\limsup_{t \rightarrow \infty} e^{-\lambda_* t} \langle g, X_t^\uparrow \rangle &= \limsup_{t \rightarrow \infty} \left(\frac{\varphi(r, v)}{p(r, v)} W_t^\uparrow - e^{-\lambda_* t} \langle \varphi/p - g, X_t^\uparrow \rangle \right) \\
&= \frac{\varphi(r, v)}{p(r, v)} W_\infty^\uparrow - \liminf_{t \rightarrow \infty} e^{-\lambda_* t} \langle \varphi/p - g, X_t^\uparrow \rangle \\
&\leq \frac{\varphi(r, v)}{p(r, v)} W_\infty^\uparrow - \langle \varphi/p - g, \tilde{\varphi} p \rangle \frac{\varphi(r, v)}{p(r, v)} W_\infty^\uparrow \\
&= \langle g, \tilde{\varphi} p \rangle W_\infty^\uparrow \frac{\varphi(r, v)}{p(r, v)},
\end{aligned}$$

as required, where we have used the normalisation $\langle \varphi, \tilde{\varphi} \rangle = 1$. □

4.4 Proof of Theorem 13

The proof we will give relies on the stochastic embedding of the skeleton process X^\uparrow in X together with a measure theoretic trick. It is worth stating the latter in the format of a proposition which is essentially taken from [15]. (The reader may note that there is a slight variation in the statement as the original version was missing an additional condition.)

Proposition 5. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ be a filtered probability space and define $\mathcal{F}_\infty := \sigma(\cup_{i=1}^\infty \mathcal{F}_i)$. Suppose $(U_t, t \geq 0)$ is an \mathcal{F} -measurable non-negative process such that $\sup_{t \geq 0} U_t$ has finite expectation and $(\mathbb{E}(U_t | \mathcal{F}_t), t \geq 0)$ is càdlàg. If*

$$\lim_{t \rightarrow \infty} \mathbb{E}(U_t | \mathcal{F}_\infty) = Y, \text{ a.s.}$$

then

$$\lim_{t \rightarrow \infty} \mathbb{E}(U_t | \mathcal{F}_t) = Y, \text{ a.s..}$$

In fact, this result can be readily obtained by considering $Y_t := \mathbb{E}(U_t | \mathcal{F}_\infty)$ then using right continuity and *Hunt's Lemma*: If $Y_n \rightarrow Y$ a.s., $(Y_n, n \in \mathbb{N})$ is dominated by $\sup_{n \in \mathbb{N}} |Y_n|$ with $\mathbb{E} \sup_{n \in \mathbb{N}} |Y_n| < \infty$, then $\mathbb{E}(Y_n | \mathcal{F}_n) \rightarrow \mathbb{E}(Y | \mathcal{F}_\infty)$ a.s..

We will take the quantities in the above proposition from their definition in the context of the physical process of the neutron transport equation. In a similar fashion to the proof of Theorem 15, set $U_t = e^{-\lambda_* t} \langle g, X_t^\uparrow \rangle$, for $g \in L_\infty^+(D \times V)$, and recall that $(\mathcal{F}_t, t \geq 0)$ is the filtration generated by the neutron branching process $(X_t, t \geq 0)$. Note that we can easily bound $(U_t, t \geq 0)$ by a multiple of $(W_t^\uparrow, t \geq 0)$ and hence we automatically get that $\sup_{t \geq 0} U_t$ has a second, and hence first, moments thanks to Corollary (3). Due to Theorem 15 and the fact that X_t^\uparrow is \mathcal{F}_∞ -measurable, $U_t = \mathbb{E}(U_t | \mathcal{F}_\infty)$ and hence

$$\lim_{t \rightarrow \infty} \mathbb{E}(U_t | \mathcal{F}_\infty) = \langle gp, \tilde{\varphi} \rangle W_\infty^\uparrow \frac{\varphi(r, v)}{p(r, v)}$$

$\mathbb{P}_{\delta_{(r, v)}}$ -almost surely, for $r \in D, v \in V$.

Using (4.65) (which comes from the skeleton embedding Theorem 14, cf. Remark 8) as we have in the proof of Theorem 15, we get

$$\mathbb{E}(U_t|\mathcal{F}_t) = \mathbb{E}(e^{-\lambda_*t}\langle g, X_t^\uparrow \rangle|\mathcal{F}_t) = e^{-\lambda_*t}\langle g, pX_t \rangle.$$

Combining this with Proposition 5 yields

$$\lim_{t \rightarrow \infty} e^{-\lambda_*t}\langle g, pX_t \rangle = \langle gp, \tilde{\varphi} \rangle W_\infty^\uparrow \frac{\varphi(r, v)}{p(r, v)}, \quad (4.88)$$

$\mathbb{P}_{\delta(r, v)}$ -almost surely. If the support of g is compactly embedded in $D \times V$, then we can replace g by g/p , with the assurance that the latter is uniformly bounded away from 0 and ∞ (cf. Lemma 10), and (4.88) gives us

$$\lim_{t \rightarrow \infty} e^{-\lambda_*t}\langle g, X_t \rangle = \langle g, \tilde{\varphi} \rangle W_\infty^\uparrow \frac{\varphi(r, v)}{p(r, v)}, \quad (4.89)$$

$\mathbb{P}_{\delta(r, v)}$ -almost surely. We can remove the assumption that the support of g is compactly embedded in $D \times V$ by appealing to similar reasoning as that of the computation in (4.87).

To complete the proof of almost sure convergence, we need to show that $W_\infty^\uparrow/p = W_\infty$, almost surely. To do so, note that if we take $g = \varphi$ in (4.89), noting that the left-hand side is equal to $\lim_{t \rightarrow \infty} W_t \varphi(r, v)$ and $\langle \varphi, \tilde{\varphi} \rangle = 1$, we get the desired result.

Finally, for the convergence in $L_2(\mathbb{P})$, first recall that we already know that $\mathbb{E}(\sup_{t \geq 0} W_t^2) < \infty$ by Doob's L_p -inequality and $L_2(\mathbb{P})$ -boundedness of W (see discussion within proof of Lemma 11). Then, by assumption $g \leq \phi$, we similarly have $\sup_{t \geq 0} \langle g, X_t \rangle$ in $L_2(\mathbb{P})$, hence we can use the dominated convergence theorem to conclude that we have convergence in $L_2(\mathbb{P})$, as well as almost surely. \square

4.5 Concluding remarks

The proof of Theorem 13 above gives a generic approach for branching particle systems which have an identified skeletal decomposition. Indeed, the reasoning is robust and will show in any such situation that the existence of a strong law of large numbers for the skeleton implies almost immediately a strong law of large numbers for the original process into which the skeleton is embedded. As an exercise, the reader is encouraged to consider the setting of a branching Brownian motion in a strip (cf. [14]). Supposing a strong law of large numbers exists on the skeleton there (in that setting it is called the 'blue tree'), then we claim that the the above reasoning applied verbatim will deliver the strong law of large numbers for the branching Brownian motion in a strip.

More generally, we claim that, modulo some minor technical modifications (e.g. taking account of the fact that E may be unbounded), in the general MBP setting of Theorem 14, an analogue of Theorem 13 may be reconstructed once the following three important components are in hand: (i) An analogue of Theorem 11; (ii) A degree of knowledge concerning the continuity properties of φ and p ; (iii) the martingale W has the property $\mathbb{E}_{\delta_x}[\sup_{t \geq 0} W_t^2] < \infty$, for all

$x \in E$. Indeed, last of these three may be weakened to γ -integrability of the martingale W , for $\gamma \in (1, 2)$, in which case one may replace many of the estimates in the Borel-Cantelli arguments by γ moment estimates instead of second moment estimates (see e.g. [13] for comparison).

It is also worth pointing out however that the reasoning in the proof of Theorem 13 does not so obviously work in the setting of superprocesses with a skeletal decomposition. Indeed a crucial step, which is automatic for branching particle systems, but less obvious for superprocesses, is the point in the argument at which we claim that $U_t = \mathbb{E}(U_t | \mathcal{F}_\infty)$. In the particle system, this statement follows immediately from the fact that \mathcal{F}_∞ carries enough information to construct the marks \uparrow and \downarrow on particles because individual genealogical lines of descent are identifiable. For superprocesses, it is less clear how to choose the filtration $(\mathcal{F}_t, t \geq 0)$ so that the notion of genealogy or otherwise can be used to claim that X_t^\uparrow , and hence U_t , is \mathcal{F}_∞ -measurable.

Acknowledgements

The body of work in this article as well as [17, 6] was born out of a surprising connection that was made at the problem formulation “Integrative Think Tank” as part of the EPSRC Centre for Doctoral Training SAMBa in the summer of 2015. We are indebted to Professor Paul Smith and Dr. Geoff Dobson from the ANSWERS modelling group at Wood for the extensive discussions as well as hosting at their offices in Dorchester. We are also grateful to Denis Villemonais for discussions on general convergence theorems for semigroups. We are also grateful to a referee and the AE who made a number of very helpful suggestions.

Bibliography

- [1] S. Asmussen and H. Hering. Strong limit theorems for general supercritical branching processes with applications to branching diffusions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36(3):195–212, 1976.
- [2] J. Berestycki, A. E. Kyprianou, and A. Murillo-Salas. The prolific backbone for supercritical superprocesses. *Stochastic Process. Appl.*, 121(6):1315–1331, 2011.
- [3] Z-Q. Chen, Y-X. Ren, R. Song, and R. Zhang. Strong law of large numbers for supercritical superprocesses under second moment condition. *Front. Math. China*, 10(4):807–838, 2015.
- [4] Z-Q. Chen, Y-X. Ren, and T. Yang. Law of large numbers for branching symmetric Hunt processes with measure-valued branching rates. *J. Theoret. Probab.*, 30(3):898–931, 2017.
- [5] A. M. G. Cox, S.C. Harris, A.E. Kyprianou, and M. Wang. Monte Carlo methods for the neutron transport equation. *Working document*, 2018.
- [6] A.M.G. Cox, S.C. Harris, E. Horton, and A.E. Kyprianou. Multi-species neutron transport equation. *J. Stat. Phys*, 176(2):425–455, 2019.
- [7] 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.
- [8] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 6*. Springer-Verlag, Berlin, 1993. Evolution problems. II, With the collaboration of Claude Bardos, Michel Cessenat, Alain Kavenoky, Patrick Lascaux, Bertrand Mercier, Olivier Pironneau, Bruno Scheurer and Rémi Sentis, Translated from the French by Alan Craig.
- [9] T. Duquesne and M. Winkel. Growth of Lévy trees. *Probab. Theory Rel. Fields*, 139(3-4):313–371, 2007.
- [10] E. B. Dynkin. *Diffusions, superdiffusions and partial differential equations*, volume 50 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2002.

- [11] E. B. Dynkin. *Superdiffusions and positive solutions of nonlinear partial differential equations*, volume 34 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2004. Appendix A by J.-F. Le Gall and Appendix B by I. E. Verbitsky.
- [12] M. Eckhoff, A. E. Kyprianou, and M. Winkel. Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Probab.*, 43(5):2545–2610, 2015.
- [13] J. Engländer, S. C. Harris, and A. E. Kyprianou. Strong law of large numbers for branching diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):279–298, 2010.
- [14] S. C. Harris, M. Hesse, and A. E. Kyprianou. Branching Brownian motion in a strip: survival near criticality. *Ann. Probab.*, 44(1):235–275, 2016.
- [15] S. C. Harris and M. I. Roberts. A strong law of large numbers for branching processes: almost sure spine events. *Electron. Commun. Probab.*, 19:no. 28, 6, 2014.
- [16] T. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)].
- [17] E. Horton, A.E. Kyprianou, and D. Villemonais. Stochastic methods for the neutron transport equation i: Linear semigroup asymptotics. *Preprint*, 2019.
- [18] E. Horton, A.E. Kyprianou, and D. Villemonais. Stochastic methods for the neutron transport equation iii: Generational many-to-one and k_{eff} . *Preprint*, 2019.
- [19] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. I. *J. Math. Kyoto Univ.*, 8:233–278, 1968.
- [20] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. II. *J. Math. Kyoto Univ.*, 8:365–410, 1968.
- [21] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes. III. *J. Math. Kyoto Univ.*, 9:95–160, 1969.
- [22] B. Lapeyre, É. Pardoux, and R. Sentis. *Introduction to Monte-Carlo methods for transport and diffusion equations*, volume 6 of *Oxford Texts in Applied and Engineering Mathematics*. Oxford University Press, Oxford, 2003. Translated from the 1998 French original by Alan Craig and Fionn Craig.
- [23] S. Maire and D. Talay. On a Monte Carlo method for neutron transport criticality computations. *IMA J. Numer. Anal.*, 26(4):657–685, 2006.
- [24] M. Mokhtar-Kharroubi. *Mathematical topics in neutron transport theory*, volume 46 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co., Inc., River Edge, NJ, 1997. New aspects, With a chapter by M. Choulli and P. Stefanov.

- [25] T. Mori, S. Watanabe, and T. Yamada. On neutron branching processes. *Publ. Res. Inst. Math. Sci.*, 7:153–179, 1971/72.
- [26] A. Murillo-Salas and J. L. Pérez. The backbone decomposition for superprocesses with non-local branching. In *XI Symposium on Probability and Stochastic Processes*, volume 69 of *Progr. Probab.*, pages 199–216. Birkhäuser/Springer, Cham, 2015.

Glossary of some commonly used notation
(Th. = Theorem, a. = above, b. = below)

Notation	Description	Introduced
$(\psi_t, t \geq 0)$	Solution to mild NTE/NBP expectation semigroup	(4.10), (4.3)
D and V	Physical and velocity domain	§4.1
σ_s, σ_f and σ	Scatter, fission and total cross-sections	b.(4.1)
π_s and π_f	Scatter and fission kernels	b.(4.1)
\overleftarrow{S} and \overleftarrow{F}	Scatter and fission operators	(4.5), (4.6)
n_{\max}	Maximum number of neutrons in a fission event	b.(4.9)
λ_*, φ and $\tilde{\varphi}$	Leading eigenvalue, right and left eigenfunctions	Th.11
$(W_t, t \geq 0)$	Additive martingale	(4.20)
<hr/>		
E and \dagger	Domain and cemetery state on which (ξ, P) is defined	§4.2.1
P and \hat{P}	Particle motion semigroup on E and $E \cup \{\dagger\}$ resp.	§4.2.1
\overleftarrow{L}	Generator associated to P for NBP	(4.61)
(ξ, P_x)	Markov process issued from $x \in E$ with P	(4.22)
(X, P_μ)	General (P, G) -MBP (and NBP)	§4.2.1
$(u_t, t \geq 0)$	Non-linear semigroup of X (and NBP)	(4.24)
ζ	Lifetime of X	(4.27)
$\varsigma(x)$	Branching rate of X at $x \in E$	§4.2.1
\mathcal{P}_x	Offspring law of X when parent at $x \in E$ (and for NBP)	a.(4.22)
G	Branching generator for MBP (and for NBP)	(4.23)
$(x_i, i = 1, \dots, N)$	Positions and number of offspring of a family in X	§4.2.1
$w(x)$	Prob. extinction when issued from $x \in E$	(4.28)
$p(x)$	Prob. survival when issued from $x \in E$	(4.58)
<hr/>		
$(X^\downarrow, P_\mu^\downarrow)$	MBP conditioned to die out	Th. 14 (i)
$(u_t^\downarrow, t \geq 0)$	Non-linear semigroup of X^\downarrow	(4.41)
P^\downarrow	Markov semigroup associated to X^\downarrow on E	Th. 14 (i)
\hat{P}^\downarrow	Markov semigroup associated to X^\downarrow on $E \cup \{\dagger\}$	Th. 14 (i)
(ξ, P_x^\downarrow)	Markov process associated to P^\downarrow	(4.35)
$\overleftarrow{L}^\downarrow$	Generator associated to P^\downarrow in the setting of NBP	(4.62)
$\varsigma^\downarrow(x)$	Branching rate of X^\downarrow at $x \in E$	(4.37)
\mathcal{P}_x^\downarrow	Offspring law of X^\downarrow	(4.38)
G^\downarrow	Branching generator of X^\downarrow	(4.36)
$(x_i^\downarrow, i = 1, \dots, N^\downarrow)$	Position and number of offspring of a family in X^\downarrow	Th. 14 (ii)

$(X^\uparrow, \mathbb{P}_\mu^\uparrow)$	Skeleton MBP (X conditioned to survive)	Th. 14 (ii)
$(X^\downarrow, \mathbb{P}_\mu^\downarrow)$	Skeleton X^\uparrow dressed with X^\downarrow trees	Th. 14 (ii)
$(u_t^\uparrow, t \geq 0)$	Non-linear semigroup of X^\downarrow	(4.51)
\mathbf{P}^\uparrow	Markov semigroup associated to X^\uparrow	Th. 14 (ii)
$(\xi, \mathbf{P}_x^\uparrow)$	Markov process associated to \mathbf{P}^\uparrow	(4.49)
$\bar{\mathbf{L}}^\uparrow$	Generator associated to \mathbf{P}^\uparrow in the setting of NBP	(4.63)
$\varsigma^\downarrow(x)$	Branching rate of X^\uparrow and X^\downarrow at $x \in E$	(4.44)
\mathcal{P}_x^\uparrow	Joint \uparrow and \downarrow offspring law of X^\downarrow when parent at $x \in E$	(4.46)
G^\uparrow	Branching generator of X^\uparrow	(4.50)
$(x^\uparrow, i = 1, \dots, N^\uparrow)$	Position and number of offspring of a family in X^\uparrow	Th. 14 (ii)
G^\downarrow	Joint branching generator of \uparrow -type and \downarrow -type in X^\downarrow	(4.45)

Concluding remarks

We considered the supercritical regime in order to characterise the growth of the system when extinction is not guaranteed. We obtained a strong law of large number result by considering a skeletal decomposition of the neutron branching process, which pertains to splitting the process into a tree which contains the particles that survive forever, and the remaining subtrees which all go extinct. For completeness, we also obtained a weak law of large numbers.

In the next chapter, we will consider a stationary eigenvalue problem. We will demonstrate the robustness of the methods developed in the previous chapters by applying them to this time-independent setting.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Stochastic Methods for Neutron Transport Equation III: Generational many-to-one and keff			
Publication status (tick one)			
Draft manuscript	<input type="checkbox"/>	Submitted	<input checked="" type="checkbox"/>
		In review	<input type="checkbox"/>
		Accepted	<input type="checkbox"/>
		Published	<input type="checkbox"/>
Publication details (reference)	Submitted to SIAM Journal on Applied Mathematics (SIAP). arXiv: 1909.00581		
Copyright status (tick the appropriate statement)			
I hold the copyright for this material	<input type="checkbox"/>	Copyright is retained by the publisher, but I have been given permission to replicate the material here	<input checked="" type="checkbox"/>
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	<p>*All research has been conducted in equal partnership with collaborators and supervisors. It is unwise to try and measure percentages of intellectual contribution as research scholarship lies as much in the escalation of ideas through mathematical discourse as it does with the seed of ideas themselves. That said, the mathematical content of this thesis is, as an entire piece of work, inextricably associated to its author through intellectual ownership.</p> <p>Formulation of ideas:</p> <p>The candidate played an integral and fully collaborative role in the formulation of ideas.</p> <p>Design of methodology:</p> <p>The candidate played an integral and fully collaborative role in the design of methodology.</p> <p>Experimental work:</p> <p>N/A</p> <p>Presentation of data in journal format:</p> <p>N/A</p> <p>*The wording in this box follows the advice and approval of my supervisor, Professor Kyprianou.</p>		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Emma Horton	Date	19/11/2019

Chapter 5

Time-independent criticality problems

Alex M. G. Cox¹, Emma Horton², Andreas E. Kyprianou³, Denis Villemonais⁴.

Remark 10. *This article has been submitted using the name Stochastic Methods for the Neutron Transport Equation III: Generational many-to-one and k_{eff} .*

Abstract

The Neutron Transport Equation (NTE) describes the flux of neutrons over time through an inhomogeneous fissile medium. In the recent articles [5, 10], a probabilistic solution of the NTE is considered in order to demonstrate a Perron-Frobenius type growth of the solution via its projection onto an associated leading eigenfunction. In [9, 4], further analysis is performed to understand the implications of this growth both in the stochastic sense, as well as from the perspective of Monte-Carlo simulation.

Such Monte-Carlo simulations are prevalent in industrial applications, in particular where regulatory checks are needed in the process of reactor core design. In that setting, however, it turns out that a different notion of growth takes centre stage, which is otherwise characterised by another eigenvalue problem. In that setting, the eigenvalue, sometimes called k -effective (written k_{eff}), has the physical interpretation as being the ratio of neutrons produced (during fission events) to the number lost (due to absorption in the reactor or leakage at the boundary) per typical fission event.

In this article, we aim to supplement [5, 10, 9, 4], by developing the stochastic analysis of the NTE further to the setting where a rigorous probabilistic interpretation of k_{eff} is given, both in terms of a Perron-Frobenius type analysis as well as via classical operator analysis.

¹Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: a.m.g.cox@bath.ac.uk

²Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: elh48@bath.ac.uk

³Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. Email: a.kyprianou@bath.ac.uk

⁴Institut Élie Cartan de Lorraine, Bureau 123, Université de Lorraine, 54506, Vandoeuvre-lès-Nancy Cedex, France. Email: denis.villemonais@univ-lorraine.fr

To our knowledge, despite the fact that an extensive engineering literature and industrial Monte-Carlo software is concentrated around the estimation of k_{eff} and its associated eigenfunction, we believe that our work is the first rigorous treatment in the probabilistic sense (which underpins some of the aforesaid Monte-Carlo simulations).

5.1 Introduction

As described in [10, 9, 5] the NTE is a balance equation for the flux of neutrons across a planar cross-section in an inhomogeneous fissile medium. The backwards form of the equation can be written as follows,

$$\begin{aligned} \frac{\partial}{\partial t} \psi_t(r, v) &= v \cdot \nabla \psi_t(r, v) - \sigma(r, v) \psi_t(r, v) \\ &+ \sigma_{\text{s}}(r, v) \int_V \psi_t(r, v') \pi_{\text{s}}(r, v, v') dv' + \sigma_{\text{f}}(r, v) \int_V \psi_t(r, v') \pi_{\text{f}}(r, v, v') dv', \end{aligned} \quad (5.1)$$

where the flux $\psi_t(r, v)$ is a function of time, t and the configuration variables $(r, v) \in D \times V$ where $D \subseteq \mathbb{R}^3$ is a non-empty, smooth, bounded convex domain such that ∂D has zero Lebesgue measure, and $V = \{v \in \mathbb{R}^3 : v_{\text{min}} \leq |v| \leq v_{\text{max}}\}$. Furthermore, the other components of (5.1) have the following interpretation:

- $\sigma_{\text{s}}(r, v)$: the rate at which scattering occurs from incoming velocity v ,
- $\sigma_{\text{f}}(r, v)$: the rate at which fission occurs from incoming velocity v ,
- $\sigma(r, v)$: the sum of the rates $\sigma_{\text{f}} + \sigma_{\text{s}}$ and is known as the total cross section,
- $\pi_{\text{s}}(r, v, v') dv'$: the scattering yield at velocity v' from incoming velocity v ,
satisfying $\int_V \pi_{\text{s}}(r, v, v') dv' = 1$, and
- $\pi_{\text{f}}(r, v, v') dv'$: the neutron yield at velocity v' from fission with incoming velocity v ,
satisfying $\int_V \pi_{\text{f}}(r, v, v') dv' < \infty$.

We also enforce the following initial and boundary conditions

$$\begin{cases} \psi_0(r, v) = g(r, v) & \text{for } r \in D, v \in V, \\ \psi_t(r, v) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } v \cdot \mathbf{n}_r > 0, \end{cases} \quad (5.2)$$

where \mathbf{n}_r is the outward unit normal at $r \in \partial D$ and $g : D \times V \rightarrow [0, \infty)$ is a bounded, measurable function. Throughout we will rely on the following assumptions in some (but not all) of our results:

(H1): Cross-sections σ_{s} , σ_{f} , π_{s} and π_{f} are uniformly bounded away from infinity.

(H2): We have $\sigma_{\text{s}} \pi_{\text{s}} + \sigma_{\text{f}} \pi_{\text{f}} > 0$ on $D \times V \times V$.

(H3): There is an open ball B compactly embedded in D such that $\sigma_{\mathbf{f}}\pi_{\mathbf{f}} > 0$ on $B \times V \times V$.

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

Note, the assumption (H1) ensures that all activity occurs at a maximum rate. Assumption (H2) ensures that at least some activity occurs, whether it be scattering or fission, together with (H3), it ensures that there is at least some fission as well as scattering. Finally (H4) is a physical constraint that is natural to nuclear fission, typically no more than 3 neutrons are produced during an average fission event. Figure 5.1 illustrates the complex nature of the inhomogeneity in the domain one typically considers.

Due to the irregular nature of gradient operator, (5.1) is meaningless in the point-wise sense, so it is often stated in one of two forms. The first is to treat (5.1) as a weak linear partial integro-differential equation (PIDE) in an appropriate Banach space, usually $L_2(D \times V)$, the space of functions $f : D \times V \mapsto [0, \infty)$ which are finite with respect to the norm $\|f\|_2 = (\int_{D \times V} f(r, v) dr dv)^{1/2}$; see e.g. [6, 7, 15]. The second is to consider the integral or *mild* form of (5.1). We refer the reader to [10, 9, 5] and the references therein for a discussion on the various formulations of the NTE and its solution. We will also elaborate on both in the forthcoming discussion.

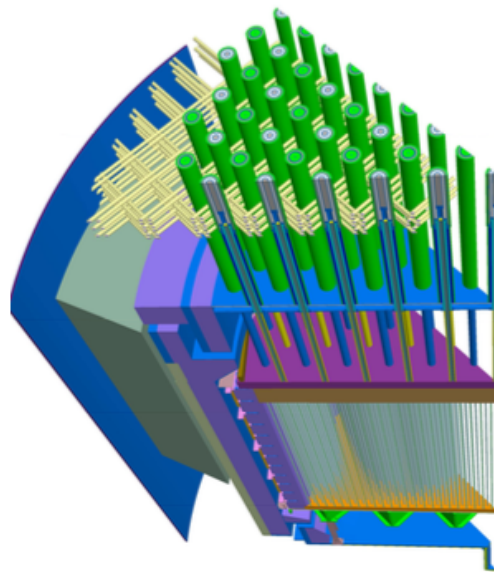


Figure 5-1: *The geometry of a nuclear reactor core representing a physical domain D , on to which the different cross-sectional values of $\sigma_{\mathbf{s}}, \sigma_{\mathbf{f}}, \pi_{\mathbf{s}}, \pi_{\mathbf{f}}$ are mapped, also as a function of neutron velocity.*

For both formats of (5.1), the papers [6, 7, 15, 10, 5] dealt with the time-eigenvalue problem and an associated Perron-Frobenius decomposition. More precisely, they give a rigorous stochastic meaning to the asymptotic

$$\psi_t \sim e^{\lambda_* t} c_g \varphi + o(e^{\lambda_* t}), \quad (5.3)$$

as $t \rightarrow \infty$, where λ_* and φ are the leading eigenvalue and associated eigenfunction associated to the NTE in the appropriate sense and c_g is a constant that depends on the initial data g .

Such an understanding is important as it promotes a number of different Monte-Carlo algorithms that can be used to estimate both the lead eigenvalue λ_* and the associated non-negative eigenfunction φ . The latter can be formulated as an eigenpair in $L_2(D \times V)$ satisfying

$$(\mathcal{T} + \mathcal{S} + \mathcal{F})\varphi = \lambda_* \varphi, \quad (5.4)$$

on $D \times V$, where

$$\begin{cases} \mathcal{T}f(r, v) & := v \cdot \nabla f(r, v) - \sigma(r, v) \\ \mathcal{S}f(r, v) & := \sigma_{\mathbf{s}}(r, v) \int_V f(r, v') \pi_{\mathbf{s}}(r, v, v') dv' \\ \mathcal{F}f(r, v) & := \sigma_{\mathbf{f}}(r, v) \int_V f(r, v') \pi_{\mathbf{f}}(r, v, v') dv', \end{cases} \quad (5.5)$$

Here, we can think of λ_* as characterising the rate of growth of flux in the system over time.

It turns out that, predominantly in industrial, engineering and (some) physics literature, there is another eigenvalue problem that plays a fundamental role in the design and safety of nuclear reactors; see for example Section 1.5 of [13]. The aforesaid eigenvalue problem involves finding (in any appropriate sense) an eigenpair k and ϕ such that

$$(\mathcal{T} + \mathcal{S})\phi + \frac{1}{k}\mathcal{F}\phi = 0. \quad (5.6)$$

The leading eigenvalue, which in the nuclear regulation industry is called k -effective, written k_{eff} , has the physical interpretation as being the ratio of neutrons produced (during fission events) to the number lost (due to absorption in the reactor or leakage at the boundary). Another interpretation of k is that it represents the average number of neutrons produced per fission event. It is this second interpretation which we exploit, since k_{eff} acts as a measure of neutrons produced between fission generations.

It is worth noting that the two eigenproblems offer potentially different sets of solutions, however, they agree in terms of criticality. More precisely, in (5.4), the triple $(\mathcal{T}, \mathcal{S}, \mathcal{F})$ is called critical if the leading eigenvalue, λ_* , in (5.4) is zero, and otherwise called subcritical (resp. supercritical) if $\lambda_* < 0$ (resp. $\lambda_* > 0$). In the setting of (5.6), the triple $(\mathcal{T}, \mathcal{S}, \mathcal{F})$ is called critical if $k_{\text{eff}} = 1$ and subcritical (resp. supercritical) if $k_{\text{eff}} < 1$ (resp. $k_{\text{eff}} > 1$).

We note however that in [2], there is a relationship between the two eigenvalues, regardless of the criticality of the system and at criticality, both (5.4) and (5.6) agree.

The main objective of this paper is to put into a rigorous setting the existence of the ‘leading’ solutions to (5.6) in the two main contexts that the NTE (5.1) is understood; that is, the weak linear PIDE context and the probabilistic context. Moreover, in the mild setting, we will build an expectation semigroup, say $(\Psi_n, n \geq 0)$, out of a stochastic process such that

$$\Psi_n[g] \sim k_{\text{eff}}^{-n} C_g \phi + o(k_{\text{eff}}^{-n}),$$

for $g \in L_{\infty}^+(D \times V)$, as $n \rightarrow \infty$, and an appropriate choice of $C_g \geq 0$. (See Theorem 22 below.) This also provides a mathematically rigorous underpinning for many of the Monte-Carlo algorithms that are used in industry for computing k_{eff} . We will offer further discussion in this direction at the end of the paper.

The rest of this article is organised as follows. In the next section, we formally introduce the description of (5.1) as a PIDE on a functional space, that is, we describe it as an abstract Cauchy problem. Moreover, we introduce two underlying stochastic processes, both of which can be used to describe the solution to the mild NTE. Also in this section, we introduce a second mild equation, (5.19), whose eigen-solutions give us a sense in which we can characterise solutions to (5.6).

In Section 5.3, we provide a solution to the newly introduced mild equation (5.19). In addition, we state the main result of this paper (Theorem 19) which shows the existence of a lead eigensolution to (5.19).

In Section 5.4, for comparison, we show how to construct and give meaning to the lead eigensolution to (5.6) in the setting of a functional space. In addition, we show how the two notions of the lead eigensolution, in this and the previous section, agree.

In Section 5.5, we give the proof of the main result of Section 5.3. Finally, we conclude in Section 5.6 with some discussion concerning the relevance of such results to previous work and Monte-Carlo methods.

5.2 Formulations of the NTE and associated eigenfunctions

As alluded to in the introduction, there are two principal ways in which the NTE is formulated. In this section, we will elaborate on them in a little more mathematical detail for later convenience and context of our main results.

5.2.1 Abstract Cauchy Problem (ACP)

Following e.g. [6, 7, 15], we want to formulate (5.1) in the function space $L_2(D \times V)$. The so-called (initial-value) abstract Cauchy problem (ACP) takes the form

$$\frac{\partial u_t}{\partial t} = \overleftarrow{\mathbf{A}} u_t \quad \text{and} \quad u_0 = g, \quad (5.7)$$

where $\overleftarrow{\mathbf{A}} = \mathcal{T} + \mathcal{S} + \mathcal{F}$ and u_t belongs to the space $L_2(D \times V)$, for $t \geq 0$ (in particular $g \in L_2(D \times V)$). Specifically, $(u_t, t \geq 0)$ is continuously differentiable in the space $L_2(D \times V)$, meaning there exists a $\dot{u}_t \in L_2(D \times V)$, which is time-continuous in $L_2(D \times V)$ with respect to $\|\cdot\|_2$ and such that $\lim_{h \rightarrow 0} h^{-1}(u_{t+h} - u_t) = \dot{u}_t$ for all $t \geq 0$. Necessarily, the solution to (5.7) forms a c_0 -semigroup⁵. Moreover, $\text{Dom}(\overleftarrow{\mathbf{A}}) := \{g \in L_2(D \times V) : v \cdot \nabla g \in L_2(D \times V) \text{ and } g|_{\partial D^+} = 0\}$ is the domain of $\overleftarrow{\mathbf{A}}$ and $u_t \in \text{Dom}(\overleftarrow{\mathbf{A}})$ for all $t \geq 0$.

⁵Recall that a c_0 -semigroup $(\mathbf{v}_t, t \geq 0)$ also goes by the name of a strongly continuous semigroup and, in the present context, this means it has the properties that (i) $\mathbf{v}_0 = \text{Id}$, (ii) $\mathbf{v}_{t+s}[g] = \mathbf{v}_t[\mathbf{v}_s[g]]$, for all $s, t \geq 0$, $g \in L_2(D \times V)$ and (iii) for all $g \in L_2(D \times V)$, $\lim_{h \rightarrow 0} \|\mathbf{v}_h[g] - g\|_2 = 0$.

Theorem 16. *Suppose (H1) holds. For $g \in L_2(D \times V)$, the unique solution to (5.7) is given by $(\mathbf{V}_t, t \geq 0)$, the c_0 -semigroup generated by $(\overleftarrow{\mathcal{A}}, \text{Dom}(\overleftarrow{\mathcal{A}}))$, i.e. the orbit $\mathbf{V}_t[g] := \exp(t\overleftarrow{\mathcal{A}})g$.*

In the ACP setting, the notion of an eigenpair (λ, φ) is well formulated on $L_2(D \times V)$ via (5.4). Equivalently, it means we are looking for $\varphi \in L_2^+(D \times V)$ and λ such that $\mathbf{V}_t[\varphi] = e^{\lambda t}\varphi$ on $L_2^+(D \times V)$, for all $t \geq 0$. The sense in which we mean that λ is a ‘leading’ eigenvalue roughly boils down to corresponding to the eigenvalue in the spectrum of the operator \mathcal{A} on $L_2(D \times V)$ with the largest real part (and, as usual, it is real valued itself), and moreover, its associated eigenfunction φ is non-negative. As such, one expects the existence of a non-negative left eigenfunction $\tilde{\varphi}$ (e.g. in the sense that $\langle \tilde{\varphi}, \mathbf{V}_t[g] \rangle = e^{\lambda t} \langle \tilde{\varphi}, g \rangle$ for $t \geq 0$) such that

$$\|e^{-\lambda t}\mathbf{V}_t[g] - \langle \tilde{\varphi}, g \rangle \varphi\|_2 = o(e^{-\lambda t}), \quad (5.8)$$

as $t \rightarrow \infty$. Here, we use the notation $\langle f, g \rangle = \int_{D \times V} f(r, v)g(r, v)drdv$, so that $\|\cdot\|_2 = \langle \cdot, \cdot \rangle^{1/2}$. Precise results of this nature can be found in [7, 15, 5].

5.2.2 Neutron branching process (NBP) and the mild NTE

We recall the neutron branching process (NBP) defined in [10], 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 $t \geq 0$. The NBP is then given by the empirical distribution of these configurations,

$$X_t(A) = \sum_{i=1}^{N_t} \delta_{(r_i(t), v_i(t))}(A), \quad A \in \mathcal{B}(D \times V), \quad t \geq 0, \quad (5.9)$$

where δ is the Dirac measure, 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 $\mathcal{M}(D \times V) := \{\sum_{i=1}^n \delta_{(r_i, v_i)} : n \in \mathbb{N}, (r_i, v_i) \in D \times V, i = 1, \dots, n\}$ which evolves randomly as follows.

A particle positioned at r with velocity v will continue to move along the trajectory $r + vt$, 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 $T_{\mathbf{s}}$ for the random time that scattering may occur, then independently of any other physical event that may affect the neutron, $\Pr(T_{\mathbf{s}} > t) = \exp\{-\int_0^t \sigma_{\mathbf{s}}(r + vs, v)ds\}$, for $t \geq 0$.

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

(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 $T_{\mathbf{f}}$ for the random time that scattering may occur, then, independently of any other physical event that may affect the neutron, $\Pr(T_{\mathbf{f}} > t) = \exp\{-\int_0^t \sigma_{\mathbf{f}}(r + vs, v) ds\}$, for $t \geq 0$.

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 $\{v_i : i = 1, \dots, N\}$. The outgoing velocities are described by the atomic random measure

$$\mathcal{Z}(A) := \sum_{i=1}^N \delta_{v_i}(A), \quad A \in \mathcal{B}(V). \quad (5.10)$$

If such an event 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 probabilities $\mathcal{P}_{(r,v)}$ are such that, for $v' \in V$, for bounded and measurable $g : V \rightarrow [0, \infty)$,

$$\int_V g(v') \pi_{\mathbf{f}}(r, v, v') dv' = \mathcal{E}_{(r,v)} \left[\int_V g(v') \mathcal{Z}(dv') \right] =: \mathcal{E}_{(r,v)}[\langle g, \mathcal{Z} \rangle]. \quad (5.11)$$

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).

The NBP is thus parameterised by the quantities $\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}$ 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 $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \mathcal{P})$ -NBP. It is associated to the NTE via the relation (5.11), and, although a $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \mathcal{P})$ -NBP is uniquely defined, a NBP specified by $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \pi_{\mathbf{f}})$ alone is not. Nonetheless, it is easy to show that for a given $\pi_{\mathbf{f}}$, a $(\sigma_{\mathbf{s}}, \pi_{\mathbf{s}}, \sigma_{\mathbf{f}}, \mathcal{P})$ -NBP always exists which satisfies (5.11). See the discussion in Section 2 of [10].

Define

$$\psi_t[g](r, v) := \mathbb{E}_{\delta_{(r,v)}}[\langle g, X_t \rangle], \quad t \geq 0, r \in \bar{D}, v \in V, \quad (5.12)$$

where $\mathbb{P}_{\delta_{(r,v)}}$ is the law of X initiated from a single particle with configuration (r, v) , and $g \in L_{\infty}^+(D \times V)$, the space of non-negative 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 (5.11)) and written $\langle g, X_t \rangle$ to mean $\int_{D \times V} g(r, v) X_t(dr, dv)$. The following result was shown in [5, 10, 7, 6].

Theorem 17. *Suppose (H1) and (H2) hold. For $g \in L_{\infty}^+(D \times V)$, the space of non-negative and uniformly bounded measurable functions on $D \times V$, there exist constants $C_1, C_2 > 0$ such that $\psi_t[g]$, as given in (5.12), 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

$$\psi_t[g] = \mathbf{U}_t[g] + \int_0^t \mathbf{U}_s[(\overleftarrow{\mathbf{S}} + \overleftarrow{\mathbf{F}})\psi_{t-s}[g]]ds, \quad t \geq 0, \quad (5.13)$$

for which (5.2) holds, where the deterministic evolution $\mathbf{U}_t[g](r, v) = g(r + vt, v)\mathbf{1}_{\{t < \kappa_{r,v}^D\}}$, $t \geq 0$, with $\kappa_{r,v}^D := \inf\{t > 0 : r + vt \notin D\}$, represents the advection semigroup associated with a single neutron travelling at velocity v from r at $t = 0$.

In [5] the below result was shown, which demonstrates the context in which the mild solution to the NTE and the ACP can be seen to coincide.

Theorem 18. *Suppose (H1) and (H2) hold. If $g \in L_\infty^+(D \times V)$ and if $(\psi_t[g], t \geq 0)$ is understood as the solution to the mild equation (5.13), then for $t \geq 0$, $\mathbf{V}_t[g] = \psi_t[g]$ on $L_2^+(D \times V)$, i.e. $\|\mathbf{V}_t[g] - \psi_t[g]\|_2 = 0$.*

In the probabilistic setting, the meaning of (5.4) can be interpreted as looking for a pair λ and φ such that, pointwise on $D \times V$, $\psi_t[\varphi] = e^{\lambda t}\varphi$, for $t \geq 0$. As alluded to in (5.3), we have a similar asymptotic to (5.8), which isolates the eigenpair (λ, φ) in its limit. The notion of ‘leading’ in the probabilistic setting is less obvious, however, due to Theorem 18, the eigenpairs that emerge from the mild setting and the weak linear PIDE setting should in principle agree on $L_2(D \times V)$. This is discussed with greater precision in [5, 10].

5.2.3 Neutron random walk (NRW)

There is a second stochastic representation of the unique solution to (5.13), which makes use of the so-called neutron random walk (NRW). This process takes values in $D \times V$ and is defined by its scatter rates, $\alpha(r, v)$, $r \in D, v \in V$, and scatter probability densities $\pi(r, v, v')$, $r \in D, v, v' \in V$. 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 $T_{\mathbf{s}}$ a scattering occurs, where $\Pr(T_{\mathbf{s}} > t) = \exp\{-\int_0^t \alpha(r + vs, v)ds\}$, for $t \geq 0$. When the scattering event occurs at position-velocity configuration (r, v) , a new velocity v' is selected with probability $\pi(r, v, v')dv'$. 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 for when it leaves the domain D , it is easy to show that (R, Υ) is a Markov process. We call the process (R, Υ) an $\alpha\pi$ -NRW.

Given a NBP defined by $\sigma_{\mathbf{s}}$, $\sigma_{\mathbf{f}}$, $\pi_{\mathbf{s}}$ and \mathcal{P} , set

$$\alpha(r, v)\pi(r, v, v') = \sigma_{\mathbf{s}}(r, v)\pi_{\mathbf{s}}(r, v, v') + \sigma_{\mathbf{f}}(r, v)\pi_{\mathbf{f}}(r, v, v') \quad r \in D, v, v' \in V. \quad (5.14)$$

and

$$\beta(r, v) = \sigma_{\mathbf{f}}(r, v) \left(\int_V \pi_{\mathbf{f}}(r, v, v')dv' - 1 \right). \quad (5.15)$$

The following result, given in [5], gives the so-called *many-to-one* representation of solution to the NTE in the form (5.13).

Lemma 12. *Suppose (H1) and (H2) hold, we have that*

$$\psi_t[g](r, v) = \mathbf{E}_{(r,v)} \left[e^{\int_0^t \beta(R_s, \Upsilon_s) ds} g(R_t, \Upsilon_t) \mathbf{1}_{\{t < \tau^D\}} \right], \quad t \geq 0, r \in D, v \in V, \quad (5.16)$$

is a second representation of the unique mild solution (in the sense of Theorem 17) of the NTE (5.13), where $\tau^D = \inf\{t > 0 : R_t \notin D\}$ and $\mathbf{P}_{(r,v)}$ for the law of the $\alpha\pi$ -NRW starting from a single neutron with configuration (r, v) .

5.2.4 Neutron generational process (NGP)

In order to solve the k -eigenvalue problem (5.6), it turns out that $(\psi_t, t \geq 0)$ and $(\phi_t, t \geq 0)$ are not the right objects to work with on account of their time-dependency. We now consider a generational model of the NBP. We can think of each line of descent in the sequence of neutron creation as genealogies. In place of $(X_t, t \geq 0)$, we consider the process $(\mathcal{X}_n, n \geq 0)$, where, for $n \geq 1$, \mathcal{X}_n is $\mathcal{M}(D \times V)$ -valued and can be written $\mathcal{X}_n = \sum_{i=1}^{\mathcal{N}_n} \delta_{(r_i^{(n)}, v_i^{(n)})}$, where $\{(r_i^{(n)}, v_i^{(n)}), i = 1, \dots, \mathcal{N}_n\}$ are the position-velocity configurations of the \mathcal{N}_n particles that are n -th in their genealogies to be the result of a fission event. \mathcal{X}_0 is consistent with X_0 and is the initial configuration of neutron positions and velocities. As such, for $n \geq 1$ we can think of \mathcal{X}_n as the n -th generation of the system and we refer to them as the neutron generational process (NGP). The reader who is more experienced with the theory of branching processes will know \mathcal{X}_n to be an example of what is called a stopping line; see [12].

Appealing to the obvious meaning of $\langle g, \mathcal{X}_n \rangle$, define the expectation semigroup $(\Psi_n, n \geq 0)$ by

$$\Psi_n[g](r, v) = \mathbb{E}_{\delta_{(r,v)}} [\langle g, \mathcal{X}_n \rangle], \quad n \geq 0, r \in D, v \in V, \quad (5.17)$$

with $\Psi_0[g] := g \in L_\infty^+(D \times V)$. The main motivation for introducing the NGP is that, just as we have seen that the meaning of (5.4) can be phrased in terms of a multiplicative invariance with respect to the solution of an ACP (5.7) or of the mild equation (5.13), we want to identify the eigen-problem (5.6) in terms of the semigroup above.

To this end, let us introduce the problem of finding a pair $k > 0$ and $\phi \in L_\infty^+(D \times V)$ such that, pointwise,

$$\Psi_1[\phi](r, v) = k\phi(r, v), \quad r \in D, v \in V. \quad (5.18)$$

In the next section we will show the existence of a solution to (5.18) which also plays an important role in the asymptotic behaviour of Ψ_n as $n \rightarrow \infty$. Before getting there, let us give a heuristic argument as to why (5.18) is another form of the eigenvalue problem (5.6).

By splitting on the first fission event, Ψ_n solves the following mild equation

$$\Psi_n[g](r, v) = \int_0^\infty \mathbf{Q}_s [\mathcal{F}\Psi_{n-1}[g]](r, v) ds, \quad r \in D, v \in V, g \in L_\infty^+(D \times V), \quad (5.19)$$

where $(\mathbb{Q}_s, s \geq 0)$ is the expectation semigroup associated with the operator $\mathcal{T} + \mathcal{S}$. More precisely,

$$\mathbb{Q}_s[g](r, v) = \mathbb{E}_{\delta_{(r,v)}} \left[e^{-\int_0^s \sigma_{\mathfrak{f}}(R_u, \Upsilon_u) du} g(R_s, \Upsilon_s) \mathbf{1}_{(s < \tau_D)} \right], \quad (5.20)$$

where $(R_s, \Upsilon_s)_{s \geq 0}$ is the $\sigma_{\mathfrak{s}}\pi_{\mathfrak{s}}$ -NRW. Then, if the pair (k, ϕ) solves (5.18), the strong Markov property along with an iteration implies that

$$k^n \phi(r, v) = \Psi_n[\phi](r, v), \quad r \in D, v \in V.$$

Using it in (5.19) and dividing through by k^n yields

$$\phi(r, v) = \int_0^\infty \mathbb{Q}_s \left[\frac{1}{k} \mathcal{F}\phi \right] (r, v) ds. \quad (5.21)$$

Now set

$$V_t := \int_0^t \mathbb{Q}_s [g] (r, v) ds.$$

Then, heuristically speaking, since \mathbb{Q} is associated to the generator $\mathcal{T} + \mathcal{S}$, classical Feynman-Kac theory suggests that V_t ‘solves’ the equation

$$\frac{\partial V_t}{\partial t} = (\mathcal{T} + \mathcal{S})V_t + g.$$

with $V_0 = 0$. Note that $\partial V_t / \partial t = \mathbb{Q}_t[g]$, which tends to zero as $t \rightarrow \infty$ thanks to the transience of (R, Υ) . Hence, taking $g = k^{-1} \mathcal{F}\phi$, letting $t \rightarrow \infty$ in the above equation, recalling that $(\mathbb{Q}_s, s \geq 0)$ is the expectation semigroup associated with the operator $\mathcal{T} + \mathcal{S}$, and using the identity (5.21) yields

$$0 = (\mathcal{T} + \mathcal{S})\phi + \frac{1}{k} \mathcal{F}\phi.$$

5.3 Probabilistic solution to (5.6)

In this section we state our main result regarding the existence of the eigenvalue and eigenfunction as specified by (5.18). We are once more motivated by the ideas presented in [3] and will use some of the techniques that were further developed in [10].

We start by constructing the many-to-one formula that is associated to the semigroup $(\Psi_n, n \geq 0)$ in the spirit of the two representations of $(\psi_t, t \geq 0)$ given in Sections 5.2.2 and 5.2.3. In this case it takes a slightly different form to the one in the time-dependent case. For ease of notation, let

$$m(r, v) := \int_V \pi_{\mathfrak{f}}(r, v, v') dv',$$

denote the mean number of offspring generated by a fission event at (r, v) , and let $(T_n, n \geq 1)$ denote the time of the scatter event in the $\alpha\pi$ -NRW that corresponds to the n -th fission event in the corresponding NBP, X .

More formally, referring to, (5.14), we can think of the $\alpha\pi$ -NRW at each scatter event as follows. For $k \geq 1$, when the NRW (R, Υ) scatters for the k -th time at (r, v) (with rate $\alpha(r, v)$), a coin is tossed and the random variable $\mathbb{I}_k(r, v)$ takes the value 1 with probability $\sigma_{\mathbf{f}}(r, v)m(r, v)/(\sigma_{\mathbf{s}}(r, v) + \sigma_{\mathbf{f}}(r, v)m(r, v))$ and its new velocity, is selected according to an independent copy of the random variable $\Theta_k^{\mathbf{f}}(r, v)$, whose distribution has probability density $\pi_{\mathbf{f}}(r, v, v')/m(r, v)$. On the other hand, with probability $\sigma_{\mathbf{s}}(r, v)/(\sigma_{\mathbf{s}}(r, v) + \sigma_{\mathbf{f}}(r, v)m(r, v))$ the random variable $\mathbb{I}_k(r, v)$ takes the value 0 and its new velocity, is selected according to an independent copy of the random variable $\Theta_k^{\mathbf{s}}(r, v)$, whose distribution has probability density $\pi_{\mathbf{s}}(r, v, v')$. As such, the velocity immediately after the n -th scatter of the NRW, given that the position-velocity configuration immediately before is (r, v) , is coded by the random variable

$$\mathbb{I}_k(r, v)\Theta_k^{\mathbf{f}}(r, v) + (1 - \mathbb{I}_k(r, v))\Theta_k^{\mathbf{s}}(r, v).$$

We thus can identify sequentially, $T_0 = 0$ and, for $n \geq 1$,

$$T_n = \inf\{t > T_{n-1} : \Upsilon_t \neq \Upsilon_{t-} \text{ and } \mathbb{I}_{k_t}(R_t, \Upsilon_{t-}) = 1\},$$

where $(k_t, t \geq 0)$ is the process counting the number of scattering events of the NRW up to time t .

Note, for the above construction of indicators to make sense, we should at least have some region of space for which fission can take place. As such the assumption (H3) becomes relevant here. Analogously to Lemma 12, we have the following many-to-one formula associated with the NBP.

Lemma 13. *Suppose (H1), (H2) and (H3) hold. The solution to (5.19) among the class of expectation semigroups is unique for $g \in L_{\infty}^+(D \times V)$ and the semigroup $(\Psi_n, n \geq 0)$ may alternatively be represented⁶ as*

$$\Psi_n[g](r, v) = \mathbf{E}_{(r, v)} \left[\prod_{i=1}^n m(R_{T_i}, \Upsilon_{T_i-}) g(R_{T_n}, \Upsilon_{T_n}) \mathbf{1}_{(T_n < \kappa^D)} \right], \quad r \in D, v \in V, n \geq 1, \quad (5.22)$$

(with $\Psi_0[g] = g$), where $(R_t, \Upsilon_t)_{t \geq 0}$ is the $\alpha\pi$ -NRW, and

$$\kappa^D := \inf\{t > 0 : R_t \notin D\}.$$

Proof. We first note that the sequence $(\Psi_n, n \geq 0)$ as defined in (5.22) is a semigroup since, due

⁶Here, we define $\prod_{i=1}^0 \cdot := 1$.

to the strong Markov property, we have

$$\begin{aligned}
& \Psi_{n+m}[g](r, v) \\
&= \mathbf{E}_{(r,v)} \left[\mathbf{E} \left[\prod_{i=1}^{n+m} m(R_{T_i}, \Upsilon_{T_i-}) g(R_{T_{n+m}}, \Upsilon_{T_{n+m}}) \mathbf{1}_{(T_{n+m} < \kappa^D)} \middle| \mathcal{F}_n \right] \right] \\
&= \mathbf{E}_{(r,v)} \left[\prod_{i=1}^n m(R_{T_i}, \Upsilon_{T_i-}) \mathbf{E}_{(R_{T_n}, \Upsilon_{T_n})} \left[\prod_{i=1}^m m(R_{T_i}, \Upsilon_{T_i-}) g(R_{T_m}, \Upsilon_{T_m}) \mathbf{1}_{(T_m < \kappa^D)} \right] \mathbf{1}_{(T_n < \kappa^D)} \right] \\
&= \Psi_n[\Psi_m[g]](r, v), \quad r \in D, v \in V.
\end{aligned}$$

In order to show that Ψ_n as defined in (5.22) does indeed solve (5.19), we consider the process at time T_1 . Before doing so, we first note that the $\alpha\pi$ -NRW is exactly the same as the $\sigma_s\pi_s$ -NRW over the time interval $[0, T_1)$ and, at time T_1 , the velocity of the former is chosen according to the expectation operator

$$\tilde{\mathcal{F}}[g](r, v) := \int_V g(r, v') \frac{\pi_{\mathfrak{f}}(r, v, v')}{m(r, v)} dv'.$$

Then, applying the strong Markov property at time T_1 ,

$$\begin{aligned}
\Psi_n[g](r, v) &= \mathbf{E}_{(r,v)} \left[\prod_{i=1}^n m(R_{T_i}, \Upsilon_{T_i-}) g(R_{T_n}, \Upsilon_{T_n}) \mathbf{1}_{(T_n < \kappa^D)} \right] \\
&= \mathbf{E}_{(r,v)} \left[m(R_{T_1}, \Upsilon_{T_1-}) \tilde{\mathcal{F}}[\Psi_{n-1}[g]](R_{T_1}, \Upsilon_{T_1-}) \mathbf{1}_{(T_1 < \kappa^D)} \right] \\
&= \int_0^\infty \mathbf{E}_{(r,v)} \left[\sigma_{\mathfrak{f}}(R_s, \Upsilon_s) e^{-\int_0^s \sigma_{\mathfrak{f}}(R_u, \Upsilon_u) du} m(R_s, \Upsilon_{s-}) \tilde{\mathcal{F}}[\Psi_{n-1}[g]](R_s, \Upsilon_{s-}) \mathbf{1}_{(s < \kappa^D)} \right] ds \\
&= \int_0^\infty \mathbb{Q}_s[\mathcal{F}\Psi_{n-1}[g]](r, v) ds,
\end{aligned}$$

where the final equality follows from the fact that $m\sigma_{\mathfrak{f}}\tilde{\mathcal{F}} = \mathcal{F}$.

It remains to show that (5.19) has a unique solution for $g \in L_\infty^+(D \times V)$ among the class of expectation semigroups, suppose that $(\Psi'_n, n \geq 0)$ is another such solution with $\Psi'_0 = g \in L_\infty^+(D \times V)$. Define $\Phi_n = \Psi_n - \Psi'_n$, for $n \geq 0$, and note by linearity that $(\Phi_n, n \geq 0)$ is another expectation semigroup with $\Phi_0 = 0$. Moreover, by linearity $(\Phi_n, n \geq 0)$ also solves (5.19). On account of this, it is straightforward to see by induction that if $\Phi_n = 0$ then $\Phi_{n+1} = 0$. The uniqueness of (5.19) in the class of expectation semigroups thus follows. \square

The next result will provide the existence of a solution to (5.18) by working directly with a variant of the semigroup $(\Psi_n, n \geq 0)$. To this end, note that, under the assumption (H4), for non-negative functions g that are bounded by one, say, we have

$$\mathbb{E}_{\delta(r,v)} [\langle g, \mathcal{X}_1 \rangle] \leq \|g\|_\infty \mathbb{E}_{\delta(r,v)} [\langle 1, \mathcal{X}_1 \rangle] \leq N_{\max}. \quad (5.23)$$

Dividing both sides of the above inequality yields a sub-Markovian semigroup. Indeed,

$$\begin{aligned}
\Psi_n^\dagger[g](r, v) &:= N_{\max}^{-n} \Psi_n[g](r, v) \\
&= \mathbf{E}_{(r,v)} \left[\prod_{i=1}^n \frac{m(R_{T_i}, \Upsilon_{T_i-})}{N_{\max}} g(R_{T_n}, \Upsilon_{T_n}) \mathbf{1}_{(T_n < \kappa^D)} \right] \\
&= \mathbf{E}_{(r,v)} \left[g(R_{T_n}, \Upsilon_{T_n}) \mathbf{1}_{(T_n < \kappa^D, n < \Gamma)} \right] \\
&=: \mathbf{E}_{(r,v)}^\dagger [g(R_{T_n}, \Upsilon_{T_n})], \tag{5.24}
\end{aligned}$$

where $\Gamma = \min\{n \geq 0 : K_n(R_{T_n}, \Upsilon_{T_n-}) = 1\}$ where, for $n \geq 0$, $r \in D$ and $v \in V$, the random variable $K_n(r, v)$ is an independent indicator random variable which is equal to 0 with probability $m(r, v)/N_{\max}$ (note, from the assumptions in Section 5.1, it follows that $\sup_{r \in D, v \in V} m(r, v) \leq N_{\max}$).

We are now ready to state the main result of this section, and indeed the article. As its proof is quite lengthy we will delay it until Section 5.5. We will need the following stronger assumption of (H3):

(H3)*: The fission cross section satisfies $\inf_{r \in D, v, v' \in V} \sigma_{\mathbf{f}}(r, v) \pi_{\mathbf{f}}(r, v, v') > 0$.

Theorem 19. *Under the assumptions (H1), (H3)* and (H4), for the semigroup $(\Psi_n, n \geq 0)$ identified by (5.19), there exist $k_* \in \mathbb{R}$, a positive⁷ right eigenfunction $\varphi \in L_\infty^+(D \times V)$ and a left eigenmeasure, η , on $D \times V$, both having associated eigenvalue k_*^n . Moreover, k_* is the leading eigenvalue in the sense that, for all $g \in L_\infty^+(D \times V)$,*

$$\langle \eta, \Psi_n[g] \rangle = k_*^n \langle \eta, g \rangle \quad (\text{resp. } \Psi_n[\varphi] = k_*^n \varphi) \quad n \geq 0, \tag{5.25}$$

and there exists $\gamma > 1$ such that, for all $g \in L_\infty^+(D \times V)$,

$$\sup_{g \in L_\infty^+(D \times V): \|g\|_\infty \leq 1} \|k_*^{-n} \varphi^{-1} \Psi_n[g] - \langle \eta, g \rangle\|_\infty = O(\gamma^{-n}) \text{ as } n \rightarrow +\infty. \tag{5.26}$$

5.4 Classical existence of solution to (5.6)

Our objective here is to make rigorous the sense in which solving (5.18) is consistent with solving the eigenvalue problem (5.6) in the classical sense.

We begin by considering the abstract Cauchy problem (ACP) on $L_2(D \times V)$,

$$\begin{cases} \frac{\partial}{\partial t} u_t &= (\mathcal{T} + \mathcal{S})u_t \\ u_0 &= g. \end{cases} \tag{5.27}$$

⁷To be precise, by a positive eigenfunction, we mean a mapping from $D \times V \rightarrow (0, \infty)$. This does not prevent it being valued zero on ∂D , as D is open.

Then, just as in the spirit of Theorems 16 and 18, it is not difficult to show that the operator $(\mathcal{T} + \mathcal{S}, \text{Dom}(\mathcal{T} + \mathcal{S}))$ generates a unique solution to (5.27) via the c_0 -semigroup $(\mathcal{V}_t, t \geq 0)$ given by

$$\mathcal{V}_t[g] := \exp(t(\mathcal{T} + \mathcal{S}))g,$$

on $L_2(D \times V)$ (and hence for $g \in L_2(D \times V)$). Moreover, we have that the expectation semigroup $(\mathcal{Q}_t, t \geq 0)$ agrees with $(\mathcal{V}_t, t \geq 0)$ on $L_2(D \times V)$, providing $g \in L_\infty^+(D \times V)$. This latter claim follows the same idea as the proof of Theorem 8.1 in [5].

The equivalence of the semigroups $(\mathcal{Q}_t, t \geq 0)$ and $(\mathcal{V}_t, t \geq 0)$ is what we will use to identify a classical (in the L_2 -sense) and probabilistic meaning to (5.6). We start by showing the classical existence of a solution to (5.6) on $L_2(D \times V)$. We note that this problem has been previously considered in [14, 15]. In [14], the author converted the criticality problem (5.6) into a time-dependent problem in order to exploit the existing theory for time-dependent problems, whereas the methods used in [15, Section 5.11] are similar to those presented in [5]. Another more restrictive version of assumption (H2) is needed, which also implies that (H3) holds:

(H5): We have $\sigma_s(r, v)\pi_s(r, v, v') > 0$ and $\sigma_f(r, v)\pi_f(r, v, v') > 0$ on $D \times V \times V$.

Theorem 20. *Suppose that the cross sections $\sigma_f\pi_f$ and $\sigma_s\pi_s$ are piecewise continuous⁸. Further, assume that (H1) and (H5) hold. Then there exist a real eigenvalue $k > 0$ and associated eigenfunction $\phi \in L_2^+(D \times V)$ such that (5.6) holds on $L_2(D \times V)$. Moreover, k can be explicitly identified as*

$$k = \sup \left\{ |\lambda| : (\mathcal{T} + \mathcal{S})\phi + \frac{1}{\lambda}\mathcal{F}\phi = 0 \text{ for some } \phi \in L_2(D \times V) \right\}. \quad (5.28)$$

Proof. We start by considering a related eigenvalue problem. First recall from [5] that, due to the transience of \mathcal{T} on D , there exist constants $M_1, \omega > 0$ such that $\|e^{t\mathcal{T}}\| \leq M_1e^{-\omega t}$ for each $t \geq 0$. Further, since \mathcal{S} is conservative, there exists $M_2 > 0$ such that⁹ $\|e^{t\mathcal{S}}\| \leq M_2, t \geq 0$. Hence $\|\mathcal{V}_t\| \leq Me^{-\omega t}, t \geq 0$, where $M = M_1M_2$. Then, classical semigroup theory [18] gives the existence of the resolvent operator $(\lambda\mathcal{I} - (\mathcal{T} + \mathcal{S}))^{-1}$ for all λ such that $\text{Re}\lambda > -\omega$, where \mathcal{I} is the identity operator on $L_2(D \times V)$. In particular, the resolvent is well-defined for $\lambda = 0$. Hence, the eigenvalue problem (5.6) is equivalent to

$$-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}\phi = k\phi. \quad (5.29)$$

Due to the assumptions (H1) and (H5), almost identical arguments to those given in the proof of [5, Proposition 9.1] show that $-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$ is a positive, compact, irreducible operator.

⁸A function is piecewise continuous if its domain can be divided into an exhaustive finite partition (e.g. polytopes) such that there is continuity in each element of the partition. This is precisely how cross sections are stored in numerical libraries for modelling of nuclear reactor cores.

⁹We use the standard definition of operator norm, namely $\|A\| = \sup_{\|f\|_2 \leq 1} \|Af\|_2$, where, as before, $\|\cdot\|_2$ is the usual norm on $L_2(D \times V)$.

Concluding in the same way as the aforementioned proposition, de Pagter's theorem [15, Theorem 5.7] implies that its spectral radius, $r(-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F})$, is strictly positive. It follows from the Krein-Rutman Theorem [5, Theorem 9.1] that $k := r(-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}) := \sup\{|\lambda| : -(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}\phi = \lambda\phi \text{ for some } \phi \in L_2(D \times V)\}$ is the leading eigenvalue of the operator $-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$ with corresponding positive eigenfunction ϕ . \square

In a similar manner to [5], we are able to provide more information about the structure of the spectrum of the operator $-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$.

Proposition 6. *Under the assumptions of Theorem 20, the part of the spectrum given by $\sigma(-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}) \cap \{\lambda : \operatorname{Re}(\lambda) > 0\}$ consists of finitely many eigenvalues with finite multiplicities. In particular, k is both algebraically and geometrically simple¹⁰.*

Proof. We follow the idea of the proof of [15, Theorem 4.13] and consider the invertibility of the operator $\lambda\mathcal{I} + (\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$ by considering the following problem,

$$\left(\mathcal{I} + \frac{1}{\lambda}(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}\right) f = \frac{1}{\lambda}g,$$

for $\lambda \in \sigma(-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}) \cap \{\lambda : \operatorname{Re}(\lambda) > 0\}$. Note that this latter set is non-empty on account of the previous theorem.

As stated in the proof of Theorem 20, the operator $-\lambda^{-1}(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$ is compact in $L_2(D \times V)$ so that by Gohberg-Shmulyan's Theorem [16], $(\mathcal{I} + \lambda^{-1}(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F})^{-1}$ exists except for a finite set of discrete degenerate poles. This implies that $(\lambda\mathcal{I} + (\mathcal{T} + \mathcal{S})^{-1}\mathcal{F})^{-1}$, $\lambda \in \sigma(-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}) \cap \{\lambda : \operatorname{Re}(\lambda) > 0\}$ exists except for a finite set of eigenvalues with finite multiplicities.

We now prove that k is a simple eigenvalue of the operator $-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$. In order to do so, we need to consider the adjoint eigenvalue problem

$$\mathcal{F}^\top(\mathcal{T}^\top + \mathcal{S}^\top)^{-1}\phi^\top = k^\top\phi^\top, \tag{5.30}$$

where \mathcal{T}^\top denotes the adjoint of \mathcal{T} , with similar definitions for \mathcal{F}^\top and \mathcal{S}^\top .

We first note that, since the operator $\mathcal{T}^\top + \mathcal{S}^\top$ enjoys similar properties to $\mathcal{T} + \mathcal{S}$, the same methods as those given in the proof of Theorem 21 apply to give existence of a leading eigenvalue k^\top and corresponding eigenfunction ϕ^\top . Now, due to [11, p.184], if λ is an isolated eigenvalue of $-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$, then its complex conjugate, $\bar{\lambda}$, is an isolated eigenvalue of the adjoint of $-(\mathcal{T} + \mathcal{S})^{-1}\mathcal{F}$ with the same multiplicity. Equivalently, for each isolated λ solving (5.6) with eigenfunction ϕ , $\bar{\lambda}$ solves (5.30) with a corresponding eigenfunction ϕ^\top and has the same multiplicity as λ . In particular, since k is real, it follows that the leading eigenvalue associated with (5.30) is also k . These observations along with similar arguments to those presented in [7, Theorem 7(iii)] and [19] yield geometric simplicity of k . Then straightforward adaptations of the arguments in [7, Remark 12] yield algebraic simplicity. \square

¹⁰An eigenvalue λ associated with an operator A is geometrically simple if $\dim(\ker(\lambda I - A)) = 1$ and algebraically simple if $\sup_{k \geq 1} \dim(\ker(\lambda I - A)^k) = 1$

The next result shows that if we can find a solution to (5.6), then it must necessarily agree with the eigensolution constructed in Theorem 19 on $L_2(D \times V)$.

Theorem 21. *Suppose the assumptions of Theorem 20 are in force¹¹, that (k_*, ϕ_*) solves (5.18) and (k, ϕ) denotes the leading eigensolution to (5.6). Then $k = k_*$, and, up to a positive constant multiple, ϕ agrees with ϕ_* on $L_2(D \times V)$.*

Proof. Recall the semigroup, $(\mathcal{V}_t)_{t \geq 0}$, generated by $\mathcal{T} + \mathcal{S}$ and note that, due to boundedness of the operator \mathcal{F} , if $g \in L_p(D \times V)$, then $\mathcal{F}g \in L_p(D \times V)$, $p \in [1, \infty]$. Thanks to [8, Chapter II, Lemma 1.3], $(\mathcal{V}_t)_{t \geq 0}$ satisfies

$$\mathcal{V}_t[\mathcal{F}g] = (\mathcal{T} + \mathcal{S}) \int_0^t \mathcal{V}_s[\mathcal{F}g] ds + \mathcal{F}g. \quad (5.31)$$

Letting $t \rightarrow \infty$ in the above equation, we obtain

$$0 = (\mathcal{T} + \mathcal{S}) \int_0^\infty \mathcal{V}_s[\mathcal{F}g] ds + \mathcal{F}g, \quad (5.32)$$

which follows from the fact that $(\mathcal{T} + \mathcal{S})$ is a transient operator so that $\lim_{t \rightarrow \infty} \mathcal{V}_t[g] = 0$. Setting $g = \phi_*$ in (5.32) and using the fact that $(\mathcal{Q}_s, s \geq 0)$ and $(\mathcal{V}_s, s \geq 0)$ agree on $L_2(D \times V)$, providing $g \in L_\infty^+(D \times V)$, yields

$$0 = (\mathcal{T} + \mathcal{S}) \int_0^\infty \mathcal{Q}_s[\mathcal{F}\phi_*] ds + \mathcal{F}\phi_*. \quad (5.33)$$

Now taking advantage of (5.18) for ϕ_* , noting in particular (5.19), we have

$$\int_0^\infty \mathcal{Q}_s[\mathcal{F}\phi_*] = \Psi_1[\phi_*] = k_*\phi_*. \quad (5.34)$$

Substituting this into (5.33) shows that (k_*, ϕ_*) is a solution to (5.6) on $L_2(D \times V)$.

To conclude the proof, we first show that $k_* = k$. Again, consider the adjoint problem (5.30) and note that

$$\begin{aligned} 0 &= \langle (\mathcal{T} + \mathcal{S})^{-1} \mathcal{F}\phi_*, \phi^\top \rangle - \langle \mathcal{F}^\top (\mathcal{T}^\top + \mathcal{S}^\top)^{-1} \phi^\top, \phi_* \rangle \\ &= (k - k_*) \langle \phi^\top, \phi_* \rangle. \end{aligned}$$

Since ϕ_* and ϕ^\top are positive, we must have $k_* = k$. Due to simplicity of k from the previous proposition, it follows that $\phi = \phi_*$ up to a multiplicative constant. \square

5.5 Proof of Theorem 19

As previously stated, our methods of proving Theorem 19 are motivated by those used in [10, 3]. The main part of the proof comes from [3, Theorem 2.1], which we restate (in the language of

¹¹Note that these assumptions imply those required for Theorem 19.

the desired application) here for convenience. To this end, recalling the notation in (5.24), define

$$\mathbf{k} = \Gamma \wedge \min\{n \geq 1 : T_n \geq \kappa^D\}.$$

Theorem 22. *Suppose that (H1), (H3)* and (H4) are in force. Suppose that there exists a probability measure ν on $D \times V$ such that*

(A1) *there exist $n_0, c_1 > 0$ such that for each $(r, v) \in D \times V$,*

$$\mathbf{P}_{(r,v)}((R_{T_{n_0}}, \Upsilon_{T_{n_0}}) \in \cdot \mid n_0 < \mathbf{k}) \geq c_1 \nu(\cdot);$$

(A2) *there exists a constant $c_2 > 0$ such that for each $(r, v) \in D \times V$ and for every $n \geq 0$,*

$$\mathbf{P}_\nu(n < \mathbf{k}) \geq c_2 \mathbf{P}_{(r,v)}(n < \mathbf{k}).$$

Then, there exists $k_c \in (0, 1)$ such that, there exist an eigenmeasure η on $D \times V$ and a positive right eigenfunction φ of Ψ_n^\dagger (defined in (5.24)) with eigenvalue k_c^n , such that η is a probability measure and $\varphi \in L_\infty^+(D \times V)$, i.e. for all $g \in L_\infty(D \times V)$,

$$\eta[\Psi_n^\dagger[g]] = k_c^n \eta[g] \quad \text{and} \quad \Psi_n^\dagger[\varphi] = k_c^n \varphi \quad n \geq 0. \quad (5.35)$$

Moreover, there exist $C, \gamma > 0$ such that, for $g \in L_\infty^+(D \times V)$ and n sufficiently large (independently of g),

$$\left\| k_c^{-n} \varphi^{-1} \Psi_n^\dagger[g] - \eta[g] \right\|_\infty \leq C \gamma^{-n} \|g\|_\infty. \quad (5.36)$$

In particular, setting $g \equiv 1$, as $n \rightarrow \infty$,

$$\left\| k_c^{-n} \varphi^{-1} \mathbf{P} \cdot (n < \mathbf{k}) - 1 \right\|_\infty \leq C \gamma^{-n}. \quad (5.37)$$

It is then straightforward to conclude that η and φ are the left eigenmeasure and right eigenfunction corresponding to the eigenvalue $k_* = k_c N_{\max}$ for the semigroup Ψ_n .

We now proceed to the proof of Theorem 22. We will use the notation J_k to denote the k^{th} scatter event of the random walk (R, Υ) under \mathbf{P}^\dagger and recall that T_k denotes the scatter event that corresponds to the k^{th} fission event in the original NBP. The basis of our proof relies on the fact that, for each $k \geq 1$, $T_k = J_k$ with positive probability.

A fundamental part of the proof of (A1) and (A2) is the following lemma. We refer the reader to [10, Lemma 7.3] for its proof.

Lemma 14. *Under the assumptions of Theorem 22, for all $r \in D$ and $v \in V$, we have*

$$\mathbf{P}_{(r,v)}^\dagger(J_7 < \mathbf{k}, R_{J_7} \in dz) \leq C \mathbf{1}_{(z \in D)} dz, \quad (5.38)$$

for some constant $C > 0$, and

$$\mathbf{P}_\nu^\dagger(J_1 < \mathbf{k}, R_{J_1} \in dz) \geq c \mathbf{1}_{(z \in D)} dz, \quad (5.39)$$

for another constant $c > 0$, where ν is Lebesgue measure on $D \times V$.

Proof of (A1). In order to prove (A1), we use similar arguments to those presented in the proof of (5.39). To this end, fix $r_0 \in D$ and suppose Υ_0 is uniformly distributed on V . Then, due to the assumptions (H1) and (H3)*, the techniques used in [10] to prove (5.39) yield

$$\mathbf{E}_{(r_0, \Upsilon_0)} [f(R_{J_1}) \mathbf{1}_{(T_1=J_1)}] \geq C_0 \int_D dz \mathbf{1}_{([r, z] \subset D)} f(z). \quad (5.40)$$

Recall the (deterministic) quantity κ_{r_0, v_0}^D , for $r_0 \in D$, $v_0 \in V$, defined in Theorem 5.13. Also note that due to (H3)*, π is bounded below by a constant (see discussion just before Lemma 7.2 of [10]). Using this, along with the strong Markov property, (H1) and (5.40), we have

$$\begin{aligned} \mathbb{E}_{(r_0, v_0)}^\dagger [f(R_{T_2}, \Upsilon_{T_2}) \mathbf{1}_{(T_2=J_2)}] &\geq C_1 \int_0^{\kappa_{r_0, v_0}^D} ds e^{-\bar{\alpha}s} \bar{\pi} \int_V dv_1 \mathbf{E}_{(r_0+v_0s, v_1)}^\dagger [f(R_{J_1}, \Upsilon_{J_1}) \mathbf{1}_{(T_1=J_1)}] \\ &\geq C_2 \kappa_{r_0, v_0}^D \int_D dr \int_V dv f(r, v). \end{aligned} \quad (5.41)$$

Finally, we note that due to (H1) and (H3)*,

$$\mathbf{P}_{(r_0, v_0)}^\dagger(T_2 < \mathbf{k}) \leq \mathbf{P}^\dagger(J_1 < \mathbf{k}) \leq \int_0^{\kappa_{r_0, v_0}^D} ds \bar{\alpha} e^{-\alpha s} \leq C_3 \kappa_{r_0, v_0}^D. \quad (5.42)$$

Combining this with (5.41) yields (A1) with ν as Lebesgue measure on $D \times V$ and $n_0 = 2$. \square

We now prove (A2). Again, we use a similar method to the one used in [10], however, we state the proof in full to illustrate where the differences occur.

Proof of A2. Let $n \geq 7$ and note that $T_n - J_7 \geq T_n - T_7$. This and the strong Markov property imply

$$\begin{aligned} \mathbf{P}_{(r, v)}(n < \mathbf{k}) &\leq \mathbf{E}_{(r, v)}^\dagger \left[\mathbf{P}_{(R_{J_7}, \Upsilon_{J_7})}(n - 7 < \mathbf{k}) \right] \\ &\leq C' \int_D \int_V \mathbf{P}_{(z, w)}(n - 7 < \mathbf{k}) dz dw, \end{aligned} \quad (5.43)$$

where we have used Lemma 14 to obtain the final inequality.

Now suppose $n \geq 1$. Recalling the measure ν from (A1), another application of Lemma 14 gives

$$\begin{aligned} \mathbf{P}_\nu(n < \mathbf{k}) &= \mathbf{E}_\nu^\dagger \left[\mathbf{1}_{(J_1 < \mathbf{k})} \mathbf{P}_{(R_{J_1}, \Upsilon_{J_1})}(n < \mathbf{k}) \right] \\ &\geq c' \int_D \int_V \mathbf{P}_{(z, w)}(n < \mathbf{k}) dz dw. \end{aligned} \quad (5.44)$$

Then, for $n \geq 8$, combining (5.43) and (5.44) yields

$$\mathbf{P}_{(r,v)}(n < \mathbf{k}) \leq \frac{C'}{c'} \mathbf{P}_\nu(n - 7 < \mathbf{k}). \quad (5.45)$$

Now recalling n_0 from (A1), it follows from (A1) that

$$\mathbf{P}_\nu^\dagger((R_{T_{n_0}}, \Upsilon_{T_{n_0}}) \in \cdot) \geq c_1 \mathbf{P}_\nu(n_0 < \mathbf{k}) \nu(\cdot). \quad (5.46)$$

Again, due to assumptions (H1) and (H3)*,

$$\mathbf{P}_\nu(n_0 < \mathbf{k}) \geq \int_{D \times V} \mathbf{P}_{(r,v)}^\dagger(T_{n_0} = J_{n_0}, n_0 < \mathbf{k}) \nu(dr, dv) \geq K, \quad (5.47)$$

for some constant $K > 0$. Then, for $n \geq 8$, due to (5.46) and (5.47),

$$\begin{aligned} \mathbf{P}_\nu(n - 7 + n_0 < \mathbf{k}) &= \mathbf{E}_\nu \left[\mathbf{1}_{(n_0 < \mathbf{k})} \mathbf{P}_{(R_{T_{n_0}}, \Upsilon_{T_{n_0}})}(n - 7 < \mathbf{k}) \right] \\ &\geq K c_1 \mathbf{P}_\nu(n - 7 < \mathbf{k}). \end{aligned} \quad (5.48)$$

Finally, noting that for $n \geq 1$ we have $n - 7 + 4n_0 \geq n$, so that

$$\mathbf{P}_\nu(n < \mathbf{k}) \geq \mathbf{P}_\nu(n - 7 + 4n_0 < \mathbf{k}),$$

and applying (5.48) four times implies

$$\mathbf{P}_\nu(n < \mathbf{k}) \geq (K c_1)^4 \mathbf{P}_\nu(n - 7 < \mathbf{k}). \quad (5.49)$$

Combining this with (5.45) yields the result. \square

5.6 Concluding remarks

We complete this paper with a number of remarks that reflect on the main theorem here and previous work in [5, 10, 9, 4].

5.6.1 λ -, k - and c -eigenvalue problems

There is a third eigenvalue problem associated with the NTE: find (c, φ_c) such that

$$\mathcal{T} \varphi_c + \frac{1}{c} (\mathcal{S} + \mathcal{F}) \varphi_c = 0.$$

The associated mild form of this eigenvalue problem is

$$\mathbf{S}_t[\varphi_c](r, v) + \frac{1}{c} \int_0^t \mathbf{S}_s[(\mathcal{S} + \mathcal{F})\varphi_c](r, v) ds = \varphi_c(r, v), \quad (5.50)$$

where

$$\mathbf{S}_t[g](r, v) = e^{-\int_0^t \sigma(r+vs, v) ds} g(r + vt, v) \mathbf{1}_{(t < \kappa_{r, v}^D)}.$$

By considering the semigroup $\Pi_n[g](r, v) = \mathbb{E}_{\delta_{(r, v)}}[\langle g, \mathbb{X}_n \rangle]$, where \mathbb{X}_n is the neutron population at the n^{th} collision (either a scatter or a fission), almost identical proofs to those given in the previous sections yield the existence of the (c, φ_c) , both in the classical sense and the probabilistic one.

In this case, the eigenvalue c can be interpreted as the ratio between neutron production (from both scattering and fission) and neutron loss (due to absorption and leakage). Alternatively, it can be seen as the number of secondary neutrons per collision, rather than only collisions due to fission events.

5.6.2 Martingale convergence and strong law of large numbers

In a similar fashion to [10], Theorem 19 implies that

$$\mathcal{W}_n := k^{-n} \frac{\langle \varphi, \mathcal{X}_n \rangle}{\langle \varphi, \mu \rangle},$$

is a non-negative mean one martingale under $\mathbb{P}_{\delta_{(r, v)}}$. One could then show that $(\mathcal{W}_n)_{n \geq 0}$ converges in $L^2(\mathbb{P})$ in the supercritical case, and otherwise has a degenerate limit.

One could also reconstruct the arguments presented in [9] to characterise the growth in the supercritical regime to obtain a strong law of large numbers:

$$\lim_{n \rightarrow \infty} k^{-n} \frac{\langle g, \mathcal{X}_n \rangle}{\langle \varphi, \mu \rangle} = \langle g, \tilde{\varphi} \rangle \mathcal{W}_\infty,$$

where \mathcal{W}_∞ is the limit of the martingale $(\mathcal{W}_n)_{n \geq 0}$.

We leave these arguments as an exercise to the reader to avoid unnecessary repetition.

5.6.3 Monte-Carlo considerations

We end this paper with a discussion of the existing Monte Carlo methods for calculating k_{eff} and the associated eigenfunctions, and how we may use the semigroup approach to propose comparable algorithms, similar in style to those presented in [4].

Due to the interpretation of the eigenvalue k_{eff} , most of the existing methods in the numerical analysis and engineering literature are based on iterative methods. For example, several algorithms are given in [17] that demonstrate how to calculate k_{eff} and φ . The main idea is to start with a set of N neutrons, distributed in $D \times V$ according to some function $\varphi^{(0)}$ that serves as an initial guess¹² at φ . The system of neutrons then evolves until the first

¹²In practice, this is usually either the uniform distribution, or the solution to a diffusion approximation of the eigenvalue problem.

generation of fission events. Letting $\hat{\varphi}^{(1)}$ be the distribution of these first generation neutrons, the first approximation, $\varphi^{(1)}$, of the eigenfunction φ is then obtained by normalising¹³ $\hat{\varphi}^{(1)}$. At the same time, the eigenvalue k_{eff} is approximated by

$$k^{(1)} = \frac{\langle \mathbf{1}, \mathcal{F}\varphi^{(1)} \rangle}{\langle \mathbf{1}, (\mathcal{T} + \mathcal{S})\varphi^{(1)} \rangle},$$

which corresponds to the ratio of source neutrons for generation 2 to the number of paths simulated in generation 1. The process is then repeated using $\varphi^{(1)}$ as the initial distribution of neutrons, in order to obtain $\varphi^{(2)}$ and $k^{(2)}$, and so on.

However, some of the methods presented in the literature lead to bias and correlations between the neutrons in successive fission generations. To overcome this problem, the notion of superhistory powering was introduced in [1]. This idea is based on letting the initial set of neutrons evolve for some number, L , of generations until the estimates for k_{eff} and φ are computed. It is usual in industry to set $L = 10$.

As we have shown in the previous sections, solving (5.6) is equivalent to look for the leading eigentriple $(k_*, \varphi, \tilde{\varphi})$ of the semigroup Ψ_n . Heuristically, this pertains to finding functions φ and $\tilde{\varphi}$ that describe where neutron production (due to fission events) is most prominent, and a parameter k_* that describes the average growth of the number of neutrons in the system. We may use the asymptotics (5.26) to inform Monte Carlo methods for the calculation of k_* , φ and $\tilde{\varphi}$. Indeed, we have

$$k_* = \lim_{n \rightarrow \infty} \frac{1}{n} \log \Psi_n[\mathbf{1}](r, v),$$

where $\mathbf{1}$ is the constant function with value one. Here, as an expectation, $\Psi_n[\mathbf{1}]$ can be approximated by Monte-Carlo simulation.

In order to calculate the eigenfunction, one can manipulate the following asymptotic.

$$\langle \tilde{\varphi}, g \rangle \varphi(r, v) = \lim_{n \rightarrow \infty} \mathbb{E}_{\delta_{(r, v)}} \left[\frac{1}{n} \sum_{m=1}^n k_*^{-m} \langle \mathcal{X}_m, g \rangle \right].$$

Varying the test function g , while keeping (r, v) fixed allows us to obtain estimates for $\tilde{\varphi}$, whereas varying the initial configuration (r, v) and keeping the test function fixed allows us to estimate φ .

Once again, the expectation can be replaced by a Monte-Carlo approximation.

We refer the reader to [4] for a more in-depth discussion of Monte Carlo algorithms based on the above asymptotics, as well as a complexity analysis of their methods. Although the algorithms and efficiency results given in [4] are for time-eigenvalues, cf. (5.4), it is straightforward to see how they may be adapted to fit the current situation (as well as their complexity). Of course, problems such as burn-in and inefficiencies that were encountered in [4] will still be present in the stationary case. We hope to carry out more formal work on this in the future.

¹³This is usually done by either setting $\varphi^{(1)} = \tilde{\varphi}^{(1)} / \|\tilde{\varphi}^{(1)}\|$ or by sampling N neutrons according to $\tilde{\varphi}^{(1)}$

Acknowledgements

We are indebted to Geoff Dobson and Paul Smith from the ANSWERS modelling group at Wood for the extensive discussions as well as hosting at their offices in Dorchester. AEK and AMGC are supported by EPSRC grant EP/P009220/1. ELH is supported by a PhD scholarship with funding from industrial partner Wood.

Bibliography

- [1] R. J. Brissenden and A. R. Garlick. Biases in the estimation of k-eff and its error by Monte Carlo methods. *Ann. nucl. Energy*, 13:63–83, 1986.
- [2] F. Brown. Fundamentals of Monte Carlo particle transport. *Lecture notes for Monte Carlo course*.
- [3] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q -process. *Probab. Theory Related Fields*, 164(1-2):243–283, 2016.
- [4] A. M. G. Cox, S.C. Harris, A.E. Kyprianou, and M. Wang. Monte Carlo methods for the neutron transport equation. *Working document*, 2018.
- [5] A.M.G. Cox, S.C. Harris, E. Horton, and A.E. Kyprianou. Multi-species neutron transport equation. *J. Stat. Phys*, 176(2):425–455, 2019.
- [6] 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.
- [7] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 6*. Springer-Verlag, Berlin, 1993. Evolution problems. II, With the collaboration of Claude Bardos, Michel Cessenat, Alain Kavenoky, Patrick Lascaux, Bertrand Mercier, Olivier Pironneau, Bruno Scheurer and Rémi Sentis, Translated from the French by Alan Craig.
- [8] K.-J. Engel and R. Nagel. *A short course on operator semigroups*. Universitext. Springer, New York, 2006.
- [9] S.C. Harris, E. Horton, and A.E. Kyprianou. Stochastic analysis of the neutron transport equation ii: Almost sure growth. *Preprint*, 2019.
- [10] E. Horton, A.E. Kyprianou, and D. Villemonais. Stochastic methods for the neutron transport equation i: Linear semigroup asymptotics. *Preprint*, 2019.
- [11] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, Berlin, 1976.
- [12] A. E. Kyprianou. Martingale convergence and the stopped branching random walk. *Probab Theory Relat. Fields*, 116(3):405–419, 2000.

- [13] E. E. Lewis and Jr. W. F. Miller. *Computational Methods of Neutron Transport*. Wiley & Sons, New York, 1984.
- [14] J. Mika. Existence and uniqueness of the solution to the critical problem in neutron transport theory. *Stud. Math.*, 37:213–225, 1971.
- [15] M. Mokhtar-Kharroubi. *Mathematical topics in neutron transport theory*, volume 46 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co., Inc., River Edge, NJ, 1997. New aspects, With a chapter by M. Choulli and P. Stefanov.
- [16] M. Ribarič and I. Vidav. Analytic properties of the inverse $a(z)^{-1}$ of an analytic linear operator valued function $a(z)$. *Comm. Pure Appl. Math.*, 11:219–242, 1958.
- [17] F. Scheben and I. G. Graham. Iterative methods for neutron transport eigenvalue problems. *SIAM J. Sci. Comput.*, 33(5):2785–2804, 2011.
- [18] J. van Neervan. *The Asymptotic Behaviour of Semigroups of Linear Operators*. Operator Theory Advances and Applications Vol. 88. Birkhäuser, 1996.
- [19] I. Vidav. Existence and uniqueness of nonnegative eigenfunctions of the Boltzmann operator. *J. Math. Anal. App.*, 22:144–155, 1968.

Concluding remarks

In this final chapter we have reconstructed the analysis used in previous chapters to prove the existence of the leading eigenvalue for the stationary problem, in both the classical L_2 setting and the expectation semigroup. In this case, we were able to show that the two settings actually agree, in the sense that the leading eigenvalues match along with their corresponding leading eigenfunction.

Since the notion of criticality in this chapter is the one most used in the nuclear industry, understanding this problem in further detail is extremely important to safety regulations and optimal reactor design. As hinted in the final discussions, the rest of the analysis carried out in previous chapters for the time-dependent problem will also apply in this setting, leading to better intuition of how to simulate the relevant quantities more efficiently.

Chapter 6

Conclusions

In this thesis, we studied criticality problems associated with the neutron transport process. Firstly, we considered the time-dependent eigenvalue problem however, as seen in the literature, there are two ways to do this, which use very different techniques. Our first step was to consolidate these two approaches. To do so, we built a branching process whose linear semigroup agreed, in an appropriate sense, with the c_0 -semigroup generated by the neutron transport integro-differential equation. We then proved the existence of the leading eigenvalue and associated left and right eigenfunction, along with Perron-Fröbenius-type decompositions, in both the analytic setting and the probabilistic setting.

Focussing on the stochastic perspective, we analysed the behaviour of the neutron branching process in the supercritical regime and obtained a spine decomposition and a skeletal decomposition. We then exploited these decompositions, in order to prove a strong law of large numbers result.

Finally, we analysed another eigenvalue problem, namely the k -eigenvalue problem. Inspired by the above analysis, we again, proved the existence of the leading eigenelements using both spectral theory and probability theory. In this case, we were able to show that the two solutions were identical.

It is worth noting that the theoretical results proved in this thesis have been used to build Monte Carlo algorithms and techniques, which we hope will complement those already in existence in industry.

We finish by discussing some of the open questions that have resulted from this research. In Chapter 2, we were able to prove the existence of the leading eigenelements, in the classical sense, for a multi-type branching process. This allowed us to capture the behaviour of the full range of particles involved in nuclear fission processes. However, in Chapter 3, when addressing the same problems in the probabilistic setting, our methods were only able to consider the movement of prompt neutrons. We hope to develop the latter methods in order to consider a multi-species model from this perspective. Another direction for future work lies in the critical regime, since this is the regime in which reactors operate in reality. We believe that it is possible to obtain a Yaglom-type limit in order to characterise the long-term behaviour of the branching process in the critical regime, conditioned on non-extinction. Finally, as previously mentioned,

we have also considered Monte Carlo algorithms in parallel to this research, however, there is still much to be done along these lines.