

YAGLOM LIMIT FOR CRITICAL NONLOCAL BRANCHING MARKOV PROCESSES

BY SIMON C. HARRIS^{1,a}, EMMA HORTON^{2,b}, ANDREAS E. KYPRIANOU^{3,c} AND MINMIN WANG^{4,d}

¹Department of Statistics, University of Auckland, ^asimon.harris@auckland.ac.nz

²Bordeaux Research Centre, INRIA, ^bemma.horton@inria.fr

³Department of Mathematical Sciences, University of Bath, ^ca.kyprianou@bath.ac.uk

⁴School of Mathematical and Physical Sciences, University of Sussex, ^dminmin.wang@sussex.ac.uk

We consider the classical Yaglom limit theorem for a branching Markov process $X = (X_t, t \geq 0)$, with nonlocal branching mechanism in the setting that the mean semigroup is critical, that is, its leading eigenvalue is zero. In particular, we show that there exists a constant $c(f)$ such that

$$\text{Law}\left(\frac{\langle f, X_t \rangle}{t} \mid \langle 1, X_t \rangle > 0\right) \rightarrow \mathbf{e}_{c(f)}, \quad t \rightarrow \infty,$$

where $\mathbf{e}_{c(f)}$ is an exponential random variable with rate $c(f)$ and the convergence is in distribution. As part of the proof, we also show that the probability of survival decays inversely proportionally to time. Although Yaglom limit theorems have recently been handled in the setting of branching Brownian motion in a bounded domain and superprocesses, (*Probab. Theory Related Fields* **173** (2019) 999–1062; *Electron. Commun. Probab.* **23** (2018) 42), these results do not allow for nonlocal branching which complicates the analysis. Our approach and the main novelty of this work is based around a precise result for the scaled asymptotics for the k th martingale moments of X (rather than the Yaglom limit itself). We then illustrate our results in the setting of neutron transport for which the nonlocality is essential, complementing recent developments in this domain (*Ann. Appl. Probab.* **30** (2020) 2573–2612; *Ann. Appl. Probab.* **30** (2020) 2815–2845; *SIAM J. Appl. Math.* **81** (2021) 982–1001; Cox et al. (2021); *J. Stat. Phys.* **176** (2019) 425–455).

1. Introduction and main results. In this article we study the long-term behaviour of a class of critical spatial Markov branching processes (MBP) on a general state space E with nonlocal branching mechanism and probabilities \mathbb{P}_{δ_x} when issued from a single particle at $x \in E$. Roughly speaking, what we mean by critical here is that the leading eigenvalue of the linear semigroup of the MBP is zero. Our objective in this article is to show that, like all other critical branching processes with finite variance offspring distributions, MBPs also respect a so-called Yaglom limit. The latter can be thought of as consisting of two components, stated precisely in Theorems 1.2 and 1.3 below, which say that for a MBP $(X_t, t \geq 0)$, for each $x \in E$ and for each bounded measurable function f on E , there exist constants $c_1(x), c(f) \in (0, \infty)$ such that

$$(1.1) \quad \lim_{t \rightarrow \infty} t \mathbb{P}_{\delta_x}(\langle 1, X_t \rangle > 0) = c_1(x),$$

and

$$(1.2) \quad \text{Law}\left(\frac{\langle f, X_t \rangle}{t} \mid \langle 1, X_t \rangle > 0\right) \text{ converges weakly to } \mathbf{e}_{c(f)},$$

Received April 2021; revised February 2022.

MSC2020 subject classifications. Primary 82D75, 60J80, 60J75; secondary 60J99.

Key words and phrases. Neutron transport equation, branching Markov process, semigroup theory, Perron-Frobenius decomposition, Yaglom limit, quasi-stationary limit.

as $t \rightarrow \infty$ where $e_{c(f)}$ is an exponentially distributed random variable with rate $c(f)$ or, equivalently, with mean $1/c(f)$.

Yaglom’s limit theorem is very classical for Galton–Watson processes, cf. [1], and a cornerstone of the theory of branching processes. Remarkably, it is only recently that Yaglom limit theorems have found their way to the literature for spatial branching processes, for example, [36] in the setting of branching Brownian motion and [37–39] in the setting of superprocesses. Unfortunately, none of the approaches taken there can be transferred to the current setting on account of the fact that there is nonlocal branching, which complicates calculations significantly. A large component of our proof of (1.1) and (1.2) (stated more precisely in Theorems 1.2 and 1.3 below) appeals to a new combinatorial methodology which is necessary on account of the spatial inhomogeneity and nonlocal nature of the branching mechanism. Specifically, our new approach and the main novelty of this work is based around an inductive argument which describes the convergence of all of the associated martingale moments of the MBP. For the proof of the latter, we make use of the so-called spine decomposition in order to extract the leading order terms using a combinatorial decomposition.

Let us now formally define our class of MBPs. Let E be a Lusin space. We will write $L^+_∞(E)$ for the space of nonnegative measurable functions, f , on E such that $\|f\| := \sup_{x \in E} f(x) < \infty$ and $L^+_{\infty,1}(E)$ for the set of functions $f \in L^+_∞(E)$ such that $\|f\| \leq 1$. A (\mathbb{P}, Φ) -MBP, $(X_t, t \geq 0)$, is an atomic measure-valued stochastic process in which particles move around independently in E according to a copy of the Markov process associated to $\mathbb{P} = (\mathbb{P}_t, t \geq 0)$. When a particle is positioned at $x \in E$, at the instantaneous rate $\gamma(x) \geq 0$, 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 the law \mathcal{P}_x . Such branching events are described by the branching mechanism Φ , defined via

$$\Phi(x, f) = \gamma(x) \mathcal{E}_x \left(\prod_{i=1}^N f(x_i) - f(x) \right), \quad f \in L^+_{\infty,1}(E), x \in E.$$

It is well known that if the configuration of particles at time t is denoted by $x_1(t), \dots, x_{N_t}(t)$, where N_t denotes the number of particles alive at time $t \geq 0$, the MBP can be described via the coordinate process in the space of atomic measures on E , denoted by $\mathcal{M}(E)$, given by

$$(1.3) \quad X_t(A) = \sum_{i=1}^{N_t} \delta_{x_i(t)}(A), \quad t \geq 0, A \in \mathcal{B}(E),$$

where δ is the Dirac measure, defined on $\mathcal{B}(E)$, the Borel subsets of E . For convenience, we will write

$$(1.4) \quad \mathcal{Z}(A) = \sum_{i=1}^N \delta_{x_i}(A), \quad A \in \mathcal{B}(E).$$

We will tend to work with the process X by describing the behaviour of $\langle f, X_t \rangle, t \geq 0$, for $f \in L^+_∞(E)$, where $\langle f, X_t \rangle := \int_E f(x) X(dx) = \sum_{i=1}^{N_t} f(x_i(t))$. Having introduced this inner product notation, we note that we will immediately abuse it by writing, for example, $\langle f, g \rangle$ to mean $\int_E f(x)g(x) dx$ as well as $\langle g, \mathcal{Z} \rangle$ to mean $\int_E g(x)\mathcal{Z}(dx) = \sum_{i=1}^N g(x_i)$.

We do not need \mathbb{P} to have the Feller property, and it is not necessary that \mathbb{P} is conservative. That said, if so desired, we can append a cemetery state $\{\dagger\}$ to E , which is to be treated as an absorbing state, and regard \mathbb{P} as conservative on the extended space $E \cup \{\dagger\}$ (which can also be treated as a Lusin space) and $\Phi(\{\dagger\}, \cdot) \equiv 0$ (i.e., no branching activity on the cemetery state). We will write \mathbb{P}_μ for the law of X with initial configuration $\mu \in \mathcal{M}(E)$ and \mathbb{E}_{δ_x} for the corresponding expectation operator.

To emphasise the generality of our model, we provide some examples of nonlocal branching processes in the literature that are encompassed by our setting. The reader is also referred to [27–29] for an introduction to general branching Markov processes.

Neutron branching process: The neutron branching process (NBP), as introduced in [11, 26], is our first example. Neutrons evolve in the configuration space $E = D \times V$, where $D \subset \mathbb{R}^3$ is a bounded, open set denoting the set of particle locations, and $V := \{v \in \mathbb{R}^3 : v_{\min} \leq |v| \leq v_{\max}\}$ with $0 < v_{\min} \leq v_{\max} < \infty$, denotes the set of velocities. From an initial space-velocity configuration (r, v) , particles move according to piecewise deterministic Markov processes, characterised by $\sigma_s \pi_s$, where $\sigma_s(r, v)$ denotes the rate at which particles change velocity (also called *scattering* events) at (r, v) , and $\pi_s(r, v, v') dv'$ denotes the probability that such a scattering event results in a new outgoing velocity v' . When at $(r, v) \in D \times V$, at rate $\sigma_f(r, v)$, a branching (or *fission*) event occurs, resulting in the release of several new neutrons with configurations $(r, v_1), \dots, (r, v_N)$, say. Thus, the nonlocality appears in the velocity variable. The nonlinear branching mechanism and associated equation was given in [20]. The NTE formed our original motivation for these results, since it is the critical regime in which nuclear reactors are operated, and thus a more complete picture of the behaviour of this setting can be obtained from Theorems 1.2 and 1.3 than results of the form (G2). Interestingly, there has been previous work on the Yaglom limit in the setting of NBPs dating back to the 1970s, cf. [33]; however, their methods relied on some rather restrictive assumptions on the NBP. We will discuss this further in Section 9.

Branching Lévy processes: Our second example is a one-dimensional branching Lévy processes, as defined in [31]. This process is defined via a triple, (Λ, L, Σ) , where a particle moves according to a Lévy process, L , and at the random time Λ which is exponentially distributed with parameter γ ; the particle dies and instantaneously scatters a random number of independent copies of X according to the point process $\Sigma(dy) = \sum_i \delta_{x_i}(dy)$ relative to its point of death.

Uchiyama process: A relative of the previous model is the one given in [41]. Particles live in \mathbb{R}^d and do not move, but after an independent and exponentially distributed random time they give birth at a constant rate γ to offspring according to an independent copy of a point process in \mathbb{R}^d centred at the parent’s position. In essence, this is a continuous-time d -dimensional branching random walk.

Biggins process: For interest, we mention the general branching process given in [2], also called the Biggins process in [30]. This process is similarly described by a triple (Z, M, χ) , where Z describes the reproduction, M describes the movement, and χ is a measure of the importance of an individual. More precisely, a particle moves according to a stochastic process, $(M(a), a \geq 0)$, where a is the age of the particle. At a random time the parent particle gives birth to a random number of individuals according to the point process Z on $\mathbb{R} \times \mathbb{R}_+$, where the first coordinate describes the spatial displacement of the child from the parent’s birth position and the second coordinate gives the parent’s age at the time of that child’s birth. Finally, χ , is another random process that can be thought of as the importance of the particle in the population as it ages. This model takes features of branching Markov processes, continuous-time branching random walks and Crump–Mode–Jagers processes. Although this process is even more general than the ones we consider in this paper, it is worth remarking that it is likely that the techniques employed here are transferable to the setting of Biggins processes, modulo some technical innovation.

Returning to the (\mathbb{P}, Φ) -branching process, recall that $(N_t, t \geq 0)$ denotes the process giving the number of particles alive at each time $t \geq 0$, and $\{x_i(t) : i = 1, \dots, N_t\}$ gives their

configurations. Then, the linear semigroup associated with X is given by

$$(1.5) \quad \psi_t[f](x) := \mathbb{E}_{\delta_x}[\langle f, X_t \rangle], \quad t \geq 0, x \in E, f \in L^+_\infty(E).$$

We now introduce our first assumption:

(G1): *The branching rate, $\gamma(x)$, and the mean number of offspring,*

$$m(x) = \mathcal{E}_x[\langle 1, \mathcal{Z} \rangle],$$

are uniformly bounded above.

We now consider the long-term average behaviour of the process X .

(G2): *There exist $\varphi \in L^+_\infty(E)$ and a finite measure $\tilde{\varphi}$ on E such that φ is uniformly bounded away from zero on each compactly embedded subset of E and for all $g \in L^+_\infty(E)$,*

$$(1.6) \quad \langle \tilde{\varphi}, \psi_t[g] \rangle = \langle \tilde{\varphi}, g \rangle \quad \text{and} \quad \psi_t[\varphi] = \varphi \quad \text{for all } t \geq 0.$$

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

$$(1.7) \quad \sup_{g \in L^+_\infty(E)} \|\varphi^{-1} \psi_t[g] - \langle \tilde{\varphi}, g \rangle\|_\infty = O(e^{-\varepsilon t}) \quad \text{for all } t \geq 0.$$

REMARK 1.1. It is usual to normalise $\tilde{\varphi}$ and φ such that $\langle \tilde{\varphi}, \varphi \rangle = 1$, which we will assume is the case throughout.

REMARK 1.2 (Criticality). Note that assumption (G2) entails that the MBP X is what we would call a ‘‘critical’’ setting. Indeed, for a lot of literature surrounding spatial branching processes, there has been emphasis on results for which an underlying assumption of exponential ergodic growth in the first moment is present; see, for example, [7, 18, 20, 26, 35, 38]. In other words, this means that the mean semigroup exhibits a Perron–Frobenius type asymptotic of the form

$$(1.8) \quad \psi_t[f](x) \sim e^{\lambda t} \varphi(x) \langle f, \tilde{\varphi} \rangle, \quad x \in E,$$

as $t \rightarrow \infty$, where λ is the lead eigenvalue of the mean semigroup ($\psi_t, t \geq 0$) and $\varphi, \tilde{\varphi}$ are the associated right eigenfunction and left eigenmeasure, respectively, meaning $\langle \mu, \psi_t[\varphi] \rangle = e^{\lambda t} \langle \mu, \varphi \rangle$ for $t \geq 0, \mu \in \mathcal{M}(E)$, and $\langle \psi_t[f], \tilde{\varphi} \rangle = e^{\lambda t} \langle \tilde{\varphi}, f \rangle$, for $f \in L^+_\infty(E), t \geq 0$, respectively.

The eigenvalue λ now gives us a sense of criticality in the spirit of the same notion for Galton–Watson processes. That is to say: The process grows on average if $\lambda > 0$ and is referred to as supercritical, decays on average if $\lambda < 0$ and referred to as subcritical and maintains a stabilising average if $\lambda = 0$ and is referred to as critical.

In this sense we see that (G2) deals with critical processes, that is, $\lambda = 0$. Although (G2) may seem like a strong assumption, it can be shown to hold for a large class of MBPs. We refer the reader to [5, 6, 26] and the references therein for various assumptions that imply (G2) as well as several examples where one can prove such results.

We will now introduce two key quantities that will appear in our main results. We first define the operator

$$(1.9) \quad \mathcal{V}[f, g](x) = \mathcal{E}_x \left[\sum_{\substack{i, j=1 \\ i \neq j}}^N f(x_i) g(x_j) \right], \quad f, g \in L^+_\infty(E), x \in E.$$

We will often abuse notation and write $\mathcal{V}[g](x)$, instead of $\mathcal{V}[g, g](x)$, for $g \in L_\infty^+(E)$ and $x \in E$. Second, we define the constant

$$(1.10) \quad \Sigma = \langle \tilde{\varphi}, \gamma \mathcal{V}[\varphi] \rangle.$$

With this notation in mind, we introduce our first main result which describes the limiting behaviour of all of the moments of the MBP and also drives the first step of the Yaglom limit.

THEOREM 1.1 (Asymptotic martingale moments). *Suppose that (G1) and (G2) hold. Further, suppose that, for some (fixed) $k \geq 1$,*

$$(1.11) \quad \sup_{x \in E} \mathcal{E}_x[\langle 1, Z \rangle^{k+1}] < \infty.$$

Then, for all $j \leq k$,

$$(1.12) \quad \sup_{x \in E} \mathbb{E}_{\delta_x}[\langle \varphi, X_t \rangle^{j+1}] < \infty, \quad t \geq 0,$$

and

$$(1.13) \quad \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \frac{1}{\varphi(x)t^j} \mathbb{E}_{\delta_x}[\langle \varphi, X_t \rangle^{j+1}] - (j+1)! \left(\frac{\Sigma}{2}\right)^j \right| = 0.$$

In order to state our other main results, we set $\zeta := \inf\{t \geq 0 : \langle 1, X_t \rangle = 0\}$ to be the extinction time of the MBP and further impose the following assumptions:

(G3): *For all $x \in E$, the process becomes extinct \mathbb{P}_{δ_x} -almost surely, that is,*

$$\mathbb{P}_{\delta_x}(\zeta < \infty) = 1.$$

(G4): *The number of offspring produced at a branching event is bounded above by a constant, N_{\max} .*

(G5): *There exists a constant $C > 0$ such that, for all $g \in L_\infty^+(E)$,*

$$\langle \tilde{\varphi}, \gamma \mathcal{V}[g] \rangle \geq C \langle \tilde{\varphi}, g \rangle^2.$$

Let us say a few words about these extra assumptions. While assumption (G2) entails that $\psi_t[g](x) \sim \langle \tilde{\varphi}, g \rangle \varphi(x)$ as $t \rightarrow \infty$, it is folklore that, for critical branching processes, the process becomes extinct almost surely, as in (G3). However, because of general spatial considerations, proving this in the general case is rather challenging. Indeed, it is possible to adapt the arguments of the proof of this result, given in [26], to yield almost sure extinction for *almost every* $x \in E$; however, one can always construct examples where the process is supercritical at a single point, for example. Thus, we include it as an assumption. We refer the reader to [19, 22–26, 40] for sufficient conditions and proofs of this result in various settings (which, e.g., include NBPs, branching diffusions and multitype continuous-time Galton–Watson processes).

Assumption (G4) is necessary for the proof of the asymptotic of the survival probability in Theorem 1.2 below. However, we emphasise that, if we assume that the conclusion of Theorem 1.2 holds, then (G4) would not be needed for the proof of Theorem 1.3; in this case we would require that all moments of the offspring distribution exist.

Assumption (G5) can be thought of as an irreducibility type condition on the branching operator, arguably playing an analogous role to, for example, the one uniform ellipticity plays for diffusion operators. Heuristically speaking, it ensures that the MBP is able to explore the state space. We will later apply this condition when g is the survival probability. In this case, the right-hand side can be seen as the survival probability of two independent particles started in the stationary distribution $\tilde{\varphi}$. Hence, (G5) implies that if there is a branching event (in the

stationary distribution), the process should have a better chance of survival than producing two particles *independently* according to $\tilde{\varphi}$.

This leads us to our main results regarding the asymptotic behaviour of the survival probability and the process conditioned to survive.

THEOREM 1.2 (Kolmogorov limit). *Suppose that assumptions (G1), (G2), (G3), (G4) and (G5) hold. Then, for all $x \in E$,*

$$\lim_{t \rightarrow \infty} t \mathbb{P}_{\delta_x}(\zeta > t) = \frac{2\varphi(x)}{\Sigma}.$$

THEOREM 1.3 (Yaglom limit). *Suppose that assumptions (G1), (G2), (G3), (G4) and (G5) hold. For $f \in L^+_{\infty}(E)$, $\theta > 0$ and for all $x \in E$,*

$$(1.14) \quad \lim_{t \rightarrow \infty} \mathbb{E}_{\delta_x} \left[\exp\left(-\theta \frac{\langle f, X_t \rangle}{t}\right) \middle| \zeta > t \right] = \frac{1}{1 + \langle \tilde{\varphi}, f \rangle \Sigma \theta / 2}.$$

As alluded to in (1.2), another way of reading Theorem 1.3 is to define $\mathbb{P}'_{\delta_x}(\cdot) = \mathbb{P}_{\delta_{(r,v)}}(\cdot | \zeta > t)$, in which case $(t^{-1} \langle f, X_t \rangle, \mathbb{P}'_{\delta_{(r,v)}})$ as $t \rightarrow \infty$ converges weakly to an exponential random variable with rate $2/\langle \tilde{\varphi}, f \rangle \Sigma$.

There does not appear to be a comparable result to Theorem 1.1 for general branching processes in the literature, least of all for processes with nonlocal branching mechanisms. However, we will make a few remarks below concerning the only branching Markov process models for which we are aware that Theorems 1.2 and 1.3 have been proved.

Isotropic neutron transport: First, let us consider [33]. Here, the authors proved Theorem 1.2 under the assumption of isotropic fission and scattering and where the fission rate and yield are independent of the position and velocity of the colliding neutron. In particular, recalling the NBP example discussed previously, we have $\sigma_{\mathbb{E}}(r, v) = \sigma$, for some constant $\sigma > 0$, and further, the law of the offspring \mathcal{P} does not depend on (r, v) .

Branching Brownian motion: Second, we consider the model given in [36]. Here, the author considers a branching Brownian motion on a compact domain D . More precisely, particles move according to a Brownian motion in D and at rate γ , a branching event occurs, producing some number of offspring at the site of the branching event. If a particle hits the boundary of D , it is killed. Again, in this case the branching rate γ and the offspring distribution are spatially independent. Furthermore, in this setting the left and right eigenfunctions are equal. Hence, the constant Σ reads

$$\Sigma = \gamma \mathcal{E}[N(N - 1)] \int_D \varphi(x)^3 dx,$$

as given in [36], Theorem 1.4.

In the spirit of the branching Brownian motion model described above, Branching Lévy processes, Uchiyama processes and Biggins processes should, in principal, be candidates from which one can similarly develop a Yaglom limit result by killing particles when they exit a compact domain in such a way that the system is critical. Generally speaking, killing in such a manner is a way to open the door to satisfying the required condition (G2).

Multitype continuous-time Galton–Watson process: Technically speaking, a multitype continuous-time Galton–Watson process is a spatial branching particle system in which E is countable. Parents do not move during their lifetime, which is independent and exponentially distributed, depending on the particle’s state say $q_i, i \in E$, and reproduce on their death, sending their offspring both to their current site as well as other sites in E in such a way that

the distribution of the offspring depends on the state of the parent. In the setting that E is finite and the mean offspring matrix satisfies an irreducibility condition, the Yaglom limit theorem is a classical result; see, for example, Chapter V.5 of [1]. In our setting the conditions of Theorem 1.3 are automatically satisfied; for example, (G2) is essentially supported by the Perron–Frobenius Theorem for the mean semigroup matrix.

For the case that E is countable but infinite, as in, for example, [15, 34], let $M(t)$ denote the matrix whose (i, j) th entry represents the mean number of type j individuals alive at time $t \geq 0$ when the process is initiated from a single type i individual and, for $\lambda \in \mathbb{R}$, set $H(\lambda) = \int_0^\infty e^{\lambda t} M(t) dt$. It was shown in [32] that if $M(t)$ is irreducible and λ is such that $H_{i,j}(\lambda) < \infty$ for a given pair $i, j \in E$, then $H_{i,j}(\lambda) < \infty$ for all $i, j \in E$. We can then define the spectral radius $\lambda_* = \sup\{\lambda \geq -\infty : H(\lambda) < \infty\}$. Thus, the first part of assumption (G2) holds, that is, (1.6); however, it is still unclear as to whether the uniform convergence in (1.7) holds in this case. Theorems 1.2 and 1.3 also require that the spectral radius satisfies $\lambda_* = 0$ but, nonetheless, give new Yaglom limit results, which we do not believe is currently in the literature.

Superprocesses: We should also note that there are recent analogues of Theorems 1.2 and 1.3 for superprocesses, cf. [37]; however, this case does not fall under our setting of branching particle systems. Moreover, the known results for superprocess are focused on local branching mechanisms. In forthcoming work we will address some of the results presented here for superprocesses with nonlocal branching mechanisms.

The rest of the paper is set out as follows. In the next section we will consider the evolution equations associated with the linear semigroup $(\psi_t, t \geq 0)$, including the many-to-one and many-to-two representations. In Section 3 we outline the main steps for the proofs of Theorems 1.1, 1.2 and 1.3, which are based around the use of the spine decomposition, together with some combinatorial arguments. Sections 4 and 5 are dedicated to the so-called spine decomposition. In particular, we provide a description this decomposition and prove a result pertaining to its ergodic behaviour, both of which will be crucial to the proof of Theorem 1.1. Sections 6, 7 and 8 contain the proofs of Theorems 1.1, 1.2 and 1.3, respectively. As alluded to earlier, it is the setting of the NTE which was the main motivation for these results in light of recent developments in this setting, cf. [10, 12, 20, 26]. Moreover, the NTE is a particular case in point where the underlying Markov process does not permit analytical formulations of the Kolmogorov and Yaglom limits which can be described in terms of PDEs. For this reason, in Section 9 we demonstrate the effectiveness of our results in the setting of the neutron branching process.

2. Evolution equations. In this section we consider some of the evolution equations associated with the linear semigroup $(\psi_t, t \geq 0)$ that will be useful in the proofs of our main results. To this end, for $f \in L_\infty^+(E)$ and $x \in E$, define

$$(2.1) \quad m[f](x) := \mathcal{E}_x[\langle f, \mathcal{Z} \rangle] = \mathcal{E}_x \left[\sum_{i=1}^N f(x_i) \right].$$

We will also use the notation $m(x) := m[\mathbf{1}](x)$ where $\mathbf{1}$ is the constant function with value 1. We further define the (linear) branching operator as follows:

$$(2.2) \quad \mathcal{A}[f](x) := \gamma(x)(m[f](x) - f(x)), \quad x \in E, f \in L_\infty^+(E).$$

Then, the evolution of $(\psi_t, t \geq 0)$ can be described in the following proposition.

PROPOSITION 2.1. *Under (G1) the semigroup $(\psi_t, t \geq 0)$ is the unique solution to the evolution equation*

$$(2.3) \quad \psi_t[f](x) = \mathbb{P}_t[f](x) + \int_0^t \mathbb{P}_s[\mathcal{A}[\psi_{t-s}[f]]](x) \, ds, \quad t \geq 0, f \in L_\infty^+(E).$$

The proof of the above proposition follows by breaking the expectation in (1.5) on the first branching event and then appealing to Lemma 1.2, Chapter 4 in [16] to tidy up the resulting integral equation. The arguments are fairly standard; see, for example, [11], and so we leave the details as an exercise to the reader.

We now consider an alternative representation of the semigroup $(\psi_t, t \geq 0)$ via a so-called many-to-one formula which will be convenient for later. Consider the process $Y = (Y_t, t \geq 0)$, which evolves according to the dynamics of $(\mathbb{P}_t, t \geq 0)$ albeit that, with instantaneous rate $\gamma(x)m(x)$, it jumps from its current position, $x \in E$, into a Borel set $A \subseteq E$ with probability $m(x)^{-1}\mathcal{E}_x[\mathcal{Z}(A)]$. Letting $\mathbf{P} = (\mathbf{P}_x, x \in E)$ denote the law of this process, we can again follow standard reasoning, highlighted above, and condition on the first first jump of Y to deduce that

$$(2.4) \quad \eta_t[f](x) := \mathbf{E}_x[f(Y_t)] = \mathbb{P}_t[f](x) + \int_0^t \mathbb{P}_s[\gamma m \hat{\mathcal{A}}[\eta_{t-s}[f]]](x) \, ds,$$

where

$$\hat{\mathcal{A}}[f](x) = \frac{1}{m(x)} \mathcal{E}_x \left[\sum_{i=1}^N f(x_i) \right] - f(x).$$

Again, in the spirit of, for example, [11], it is not difficult to show that one can relate the linear semigroup $(\psi_t, t \geq 0)$ to the process Y via the following many-to-one formula.

LEMMA 2.1 (Many-to-one formula). *Suppose (G1) is satisfied. Then,*

$$(2.5) \quad \psi_t[f](x) = \mathbf{E}_x[e^{\int_0^t B(Y_s) \, ds} f(Y_t)], \quad t \geq 0, f \in L_\infty^+(E),$$

where $B(x) = \gamma(x)(m(x) - 1)$.

In a similar manner, we can consider a many-to-two formula. For this, recall the operator \mathcal{V} , introduced in (1.9).

LEMMA 2.2. *Let $g \in L_\infty^+(E)$, $x \in E$ and $t \geq 0$. Suppose (G1) holds and that*

$$\sup_{x \in E} \mathcal{E}_x[\langle 1, \mathcal{Z} \rangle^2] < \infty.$$

Then, we have the following many-to-two formula:

$$(2.6) \quad \mathbb{E}_{\delta_x}[\langle f, X_t \rangle \langle g, X_t \rangle] = \psi_t[fg](x) + \int_0^t \psi_s[\gamma \mathcal{V}[\psi_{t-s}[f], \psi_{t-s}[g]]](x) \, ds.$$

As with the proof of Proposition 2.1, we have excluded the proofs of Lemma 2.1 and Lemma 2.2, as they are fairly standard. In addition to [11], we refer the reader to [12, 26] for the proofs in the case of the NBP, discussed in the Introduction, and to [21] for some general theory in the case of local branching.

3. Three step approach to the Yaglom limit. The way we will approach the proofs of Theorems 1.1, 1.2 and 1.3 is through three fundamental steps.

3.1. *Step 1.* The first step, which is detailed in Section 6, deals with the limiting behaviour of the martingale moments of the MBP, as presented in Theorem 1.1. To this end, let us consider the following martingale change of measure:

$$(3.1) \quad \frac{d\mathbb{P}_\mu^\varphi}{d\mathbb{P}_\mu} \Big|_{\mathcal{F}_t} = \frac{\langle \varphi, X_t \rangle}{\langle \varphi, \mu \rangle}, \quad t \geq 0, \mu \in \mathcal{M}(E).$$

Note that this change of measure is well defined, due to the fact that, under (G2), φ is an eigenfunction, that is,

$$(3.2) \quad \mathbb{E}_\mu[\langle \varphi, X_t \rangle] = \langle \varphi, \mu \rangle, \quad t \geq 0, \mu \in \mathcal{M}(E),$$

together with the branching Markov property.

Via a combination of the spine decomposition induced by (3.1) (which we will discuss in the next section), along with general results for Poisson processes and some combinatorial results, we will prove the following result.

THEOREM 3.1. *Under the assumptions of Theorem 1.1 for some fixed $k \geq 1$, we have, for all $j \leq k$,*

$$(3.3) \quad \sup_{x \in E} \mathbb{E}_{\delta_x}^\varphi [\langle \varphi, X_t \rangle^j] < \infty, \quad t \geq 0$$

and

$$(3.4) \quad \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \frac{1}{t^j} \mathbb{E}_{\delta_x}^\varphi [\langle \varphi, X_t \rangle^j] - (j + 1)! \left(\frac{\Sigma}{2} \right)^j \right| = 0.$$

Note that, thanks to (3.1), this is equivalent to Theorem 1.1. This result is new in its own right, independently of the Yaglom limit, and little seems to be known in the literature for (asymptotic) limits of spatial branching processes.

3.2. *Step 2.* For the next step, which we handle in Section 7, we prove Theorem 1.2. Our approach will take inspiration from the original approach to the Yaglom limit theorem for the neutron branching process in Section 5 of [33]. As such, we develop a nonlinear integral equation for the survival probability $\mathbb{P}_{\delta_x}(\zeta > t)$ and show that, using an analytical approach of the operators involved, its asymptotic behaviour is linear. However, one major difference with the setting of Section 5 of [33] is that the generality of the branching mechanism in this setting forces a significantly more complex analysis of the aforementioned nonlinear equation. Hence, we first obtain coarse upper and lower bounds of order $1/t$ which are then used to obtain the precise limit.

We note that combining steps 1 and 2 then concludes the proof of Theorem 1.3, but only in the special case when $f = \varphi$. Indeed,

$$(3.5) \quad \begin{aligned} \mathbb{E}_{\delta_x}[\langle \varphi, X_t \rangle^k | \zeta > t] &= \frac{1}{t^k} \frac{\mathbb{E}_{\delta_x}[\langle \varphi, X_t \rangle^k \mathbf{1}_{\{\zeta > t\}}]}{\mathbb{P}_{\delta_x}(\zeta > t)} \\ &= \varphi(x) \frac{\mathbb{E}_{\delta_x}^\varphi[\langle \varphi, X_t \rangle^{k-1} / t^{k-1}]}{t \mathbb{P}_{\delta_x}(\zeta > t)} \\ &\rightarrow k! \left(\frac{\Sigma}{2} \right)^k, \quad t \rightarrow \infty \end{aligned}$$

which gives the desired result on account of the fact that the exponential distribution is uniquely identifiable by its moments. We refer the reader to [3], Section 30, and [4], Theorem 8.48, for more general results regarding classical moment convergence problems.

3.3. *Step 3.* At this point in the argument, we have established the desired convergence of martingale moments and hence the Yaglom limit (1.14) for the case that $f = \varphi$. In the final step, detailed in Section 8, we remove this restriction by considering the decomposition

$$f = \tilde{f} + \langle \tilde{\varphi}, f \rangle \varphi \quad \text{where } \tilde{f} = f - \langle \tilde{\varphi}, f \rangle \varphi.$$

The reasoning is a relatively standard way to wrap things up, for example, as in [36, 37, 39]. We use a second moment estimate to show that, for all $x \in E$, $\langle \tilde{f}, X_t \rangle / t$ converges weakly to zero under $\mathbb{P}_{\delta_x}(\cdot | \zeta > t)$, as $t \rightarrow \infty$. As such, the principal contribution in the convergence of $\langle f, X_t \rangle / t$ as $t \rightarrow \infty$ is its projection on to φ , that is, the limit of $\langle \tilde{\varphi}, f \rangle \langle \varphi, X_t \rangle / t$.

4. Spine decomposition. Recall that

$$W_t := \frac{\langle \varphi, X_t \rangle}{\langle \mu, \varphi \rangle}, \quad x \in E, t \geq 0,$$

is a mean one martingale and can thus be used to define the following change of measure:

$$(4.1) \quad \frac{d\mathbb{P}_\mu^\varphi}{d\mathbb{P}_\mu} \Big|_{\mathcal{F}_t} = \frac{\langle \varphi, X_t \rangle}{\langle \varphi, \mu \rangle}, \quad t \geq 0, \mu \in \mathcal{M}(E),$$

for the (\mathbb{P}, Φ) -MBP. It is well known that, for many branching processes, the change of measure (4.1) induces the so-called spine decomposition. Since this will be key to the proof of Theorem 1.1, we first spend some time discussing it.

To this end, let us introduce another $\mathcal{M}(E)$ -valued stochastic process via the following pathwise construction:

1. From the initial configuration $\mu \in \mathcal{M}(E)$ with an arbitrary enumeration of particles, the i^* th individual is selected and marked “*spine*” so that, given μ ,

$$(4.2) \quad \text{the probability } i^* = j \text{ is given by } \frac{\varphi(x_j)}{\langle \varphi, \mu \rangle} \text{ for } j = 1, \dots, N.$$

2. The individuals $j \neq i^*$ in the initial configuration that are not marked “*spine*,” each issue independent copies of $(X, \mathbb{P}_{\delta_{x_j}})$ respectively.

3. The marked individual issues a single particle whose motion is determined by the semi-group

$$\mathbb{P}_t^\varphi[f](x) := \frac{1}{\varphi(x)} \mathbf{E}_x \left[e^{\int_0^t \frac{B(Y_s)}{\varphi(Y_s)} (m[\varphi](Y_s) - \varphi(Y_s)) ds} \varphi(Y_t) g(Y_t) \right], \quad x \in E, f \in L_\infty^+(E).$$

4. The marked individual undergoes branching at the accelerated rate

$$(4.3) \quad \rho(x) := \gamma(x) \frac{m[\varphi](x)}{\varphi(x)}$$

when in physical configuration $x \in E$, at which point it scatters a random number of particles according to the random measure on E , given by $(\mathcal{Z}, \mathcal{P}_x^\varphi)$, where

$$(4.4) \quad \frac{d\mathcal{P}_x^\varphi}{d\mathcal{P}_x} = \frac{\langle \varphi, \mathcal{Z} \rangle}{m[\varphi](x)}.$$

5. Given \mathcal{Z} from the previous step, μ is redefined as $\mu = \mathcal{Z}$, and step 1 is repeated.

The process above describes the physical configuration of all the particles in the system at time $t \geq 0$, that is, ignoring the marked genealogy. The spine decomposition, induced by (3.1), is summarised by the following theorem.

THEOREM 4.1. *Under assumptions (G1) and (G2), the process constructed above in 1–5 is Markovian and equal in law to (X, \mathbb{P}^φ) , where $\mathbb{P}^\varphi = (\mathbb{P}_\mu^\varphi, \mu \in \mathcal{M}(E))$ was defined in (4.1).*

We are also be interested in the evolution of the single marked genealogical line of descent which will be referred to simply as *the spine*. The following lemma identifies the spine as a change of measure applied to the process Y defined in Lemma 2.1.

LEMMA 4.1. *Under assumptions (G1) and (G2), the spine is equal in law to (Y, \mathbf{P}^φ) , where*

$$(4.5) \quad \frac{d\mathbf{P}_x^\varphi}{d\mathbf{P}_x} \Big|_{\mathcal{F}_t} = e^{\int_0^t B(Y_s) ds} \frac{\varphi(Y_t)}{\varphi(x)} \mathbf{1}_{\{t < \kappa\}}, \quad t \geq 0, x \in E,$$

where $\kappa := \inf\{t \geq 0 : Y_t \notin E\}$, and we recall that $B(x) = \gamma(x)(m(x) - 1)$. From this we see that the semigroup $(\mathbb{P}_t^\varphi, t \geq 0)$, associated to (Y, \mathbf{P}^φ) , is conservative and satisfies, for $f \in L_\infty^+(E)$, $\mathbb{E}_t^\varphi[f] = \varphi^{-1} \psi_t[\varphi f]$, $t \geq 0$ with stationary limit distribution

$$\varphi(x) \tilde{\varphi}(dx), \quad x \in E.$$

In particular, the above two results mean that under \mathbb{E}^φ , we can write

$$(4.6) \quad \frac{\langle \varphi, X_t \rangle}{t} = \frac{\varphi(Y_t)}{t} + \frac{1}{t} \sum_{j=1}^{n_t} \Xi_j(Y_{T_j-}, t - T_j),$$

where n_t is the number of branching events along the spine up to time t , whose times are denoted $(T_j, j \geq 1)$ and, given $x \in E$ and $0 \leq u < \infty$, the random variables $\Xi_j(x, u)$ are independent and equal in law to

$$(4.7) \quad \sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_u^i \rangle \quad \text{under } \eta_x^\varphi := \mathcal{P}_x^\varphi \bigotimes_{i=1}^N \mathbb{P}_{\delta_{x_i}},$$

where i^* denotes the index of the spine particle. To be more precise, η_x^φ is the law of producing offspring, according to \mathcal{P}_x^φ , from which the spine index i^* can be selected empirically, according to (4.2), and then $\langle \varphi, X_u^i \rangle$ can be sampled for $i \neq i^*$ under the respective $\mathbb{P}_{\delta_{x_i}}$. Under \mathbb{P}^φ , particles alive at time t thus consist of the spine particle at time t plus those that come from independent copies of the original process that have branching from the spine before time t . The latter are described by the variables Ξ_j , which are correlated with the path of Y , and occur at the arrival times of n_t at rate

$$(4.8) \quad \rho(Y_{t-}) \eta_{Y_{t-}}^\varphi(d\Xi) dt.$$

Note also that, as $\|\varphi\|_\infty < \infty$, the spine term in (4.6) converges uniformly to zero.

5. Ergodic properties of the spine. Next, we state and prove a theorem relating to the long-term behaviour of the spine which will be of general use throughout our computations. It is loosely based on Lemma 5.1 of [39]. We emphasise that it only concerns the spine and not the whole branching process; thus, we will only need to work with (Y, \mathbf{P}^φ) . In order to state it, let us introduce a class of functions \mathcal{C} on $E \times [0, \infty) \times [0, \infty)$ such that F belongs to class \mathcal{C} if

$$F(x, s) := \lim_{t \rightarrow \infty} F(x, s, t), \quad x \in E, s \geq 0,$$

exists,

$$(5.1) \quad \sup_{x \in E, 0 < s < 1} |\varphi(x)F(x, s)| < \infty,$$

and

$$(5.2) \quad \lim_{t \rightarrow \infty} \sup_{x \in E} \varphi(x) |F(x, s) - F(x, s, t)| = 0.$$

THEOREM 5.1. *Suppose (G1) and (G2) hold. For any $k \geq 1$ and $F_1, \dots, F_k \in \mathcal{C}$, we have*

$$\lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{i=1}^k \int_0^1 F_i(Y_{ut}, u, t) \, du \right] - \prod_{i=1}^k \int_0^1 \langle \varphi \tilde{\varphi}, F_i(\cdot, u) \rangle \, du \right| = 0,$$

where $F_i(\cdot, u) = \lim_{t \rightarrow \infty} F_i(\cdot, u, t)$.

PROOF. We will show that, for $0 \leq u_1, \dots, u_k \leq 1$ and $F_1, \dots, F_k \in \mathcal{C}$,

$$(5.3) \quad \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{i=1}^k F_i(Y_{u_i t}, u_i, t) \right] - \prod_{i=1}^k \langle \varphi \tilde{\varphi}, F_i(\cdot, u_i) \rangle \right| = 0.$$

Once that has been established, it is a simple matter of dominated convergence to deduce that

$$\begin{aligned} & \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{i=1}^k \int_0^1 F_i(Y_{ut}, u, t) \, du \right] - \prod_{i=1}^k \int_0^1 \langle \varphi \tilde{\varphi}, F_i(\cdot, u) \rangle \, du \right| \\ &= \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\int_0^1 \cdots \int_0^1 \prod_{i=1}^k F_i(Y_{u_i t}, u_i, t) \, du_1 \cdots du_k \right] \right. \\ & \quad \left. - \int_0^1 \cdots \int_0^1 \prod_{i=1}^k \langle \varphi \tilde{\varphi}, F_i(\cdot, u_i) \rangle \, du_1 \cdots du_k \right| \\ &\leq \int_0^1 \cdots \int_0^1 \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{i=1}^k F_i(Y_{u_i t}, u_i, t) \right] - \prod_{i=1}^k \langle \varphi \tilde{\varphi}, F_i(\cdot, u_i) \rangle \right| \, du_1 \cdots du_k \\ &= 0, \end{aligned}$$

as required.

The proof of (5.3) works by induction. We start by proving (5.3) in the case that $k = 1$. Suppose $F \in \mathcal{C}$. As $\mathbf{E}_x^\varphi [F(Y_{ut}, u, t)] = \mathbb{P}_{ut}^\varphi [F(\cdot, u, t)](x)$, we can use Lemma 4.1 and write

$$(5.4) \quad \begin{aligned} & \sup_{x \in E} |\mathbb{P}_{ut}^\varphi [F(\cdot, u, t)](x) - \langle \varphi \tilde{\varphi}, F(\cdot, u) \rangle| \\ & \leq \sup_{x \in E} \frac{1}{\varphi(x)} |\psi_{ut}[\varphi F(\cdot, u, t)](x) - \psi_{ut}[\varphi F(\cdot, u)](x)| \\ & \quad + \sup_{x \in E} \left| \frac{1}{\varphi(x)} \psi_{ut}[\varphi F(\cdot, u)](x) - \langle \tilde{\varphi}, \varphi F(\cdot, u) \rangle \right|. \end{aligned}$$

Since ψ_t is a linear operator, from (5.2), for t sufficiently large, the first term on the right-hand side of (5.4) can be estimated by $\varphi^{-1} \psi_{ut}[\varepsilon]$ for each $\varepsilon > 0$. Hence, by the triangle inequality

the fact that $\langle \tilde{\varphi}, \varepsilon \rangle \leq c\varepsilon \langle \tilde{\varphi}, 1 \rangle$ for some constant $c > 0$ and (G2), for t sufficiently large,

$$\begin{aligned}
 & \sup_{x \in E} |\mathbb{P}_{ut}^\varphi[F(\cdot, u, t)](x) - \langle \varphi \tilde{\varphi}, F(\cdot, u) \rangle| \\
 (5.5) \quad & \leq \sup_{x \in E} \left| \frac{1}{\varphi(x)} \psi_{ut}[\varepsilon](x) - \langle \tilde{\varphi}, \varepsilon \rangle \right| + \varepsilon \|\tilde{\varphi}\|_1 \\
 & \quad + \sup_{x \in E} \left| \frac{1}{\varphi(x)} \psi_{ut}[\varphi F(\cdot, u)](x) - \langle \tilde{\varphi}, \varphi F(\cdot, u) \rangle \right| \\
 & \leq C(e^{-\alpha t} + \varepsilon)
 \end{aligned}$$

for some strictly positive constants C, α which exist thanks to (G2). This ensures that the left-hand side of (5.4) tends to zero as $t \rightarrow \infty$, as required.

Now, suppose that (5.3) is true for $k = n - 1$ and, without loss of generality, that $0 \leq u_1 \leq \dots \leq u_n \leq 1$. Then, by the Markov property

$$\mathbb{E}_x^\varphi \left[\prod_{i=1}^n F_i(Y_{u_i t}, u_i, t) \right] = \int_E \mathbf{P}_x^\varphi(Y_{u_1 t} \in dy') F_1(y', u_1, t) \mathbb{E}_{y'}^\varphi \left[\prod_{i=2}^n F_i(Y_{(u_i - u_1)t}, u_i, t) \right].$$

Now, define

$$\tilde{F}(y', u_1, t) = F_1(y', u_1, t) \mathbb{E}_{y'}^\varphi \left[\prod_{i=2}^n F_i(Y_{(u_i - u_1)t}, u_i, t) \right].$$

Note, moreover, that, using the induction hypothesis,

$$\tilde{F}(y', u_1) := \lim_{t \rightarrow \infty} \tilde{F}(y', u_1, t) = F_1(y', u_1) \prod_{i=2}^n \langle \varphi \tilde{\varphi}, F_i(\cdot, u_i) \rangle, \quad y' \in E,$$

which further respects the required boundedness conditions in the statement of the proposition. Hence, appealing again to the induction hypothesis for $n = 1$, we get the required result. \square

6. Step 1: Proof of Theorem 1.1. We now prove Theorem 1.1 and thus assume that (G1), (G2) and (1.11) are in force. The strategy of the proof will be to use induction. For the inductive step we use an interim calculation that looks at the asymptotic behaviour of contributions to the k th moment that comes from the subtrees along the spine.

PROOF. Recall that Theorem 1.1 can be equivalently formulated as Theorem 3.1. As previously mentioned, the spine decomposition will play an important role in the proof, and we will rely heavily on the decomposition (4.6). In addition to this, a number of combinatorial results for moment formulae will come into play for which we have placed two fundamental results in the Appendix that we will call upon. The proof is long, and we break it into a number of blocks of calculations which fall under informal headings:

Block 1: The theorem holds for $k = 1$.

We start by noting that, the moment assumption (1.11) implies that both $\sup_{x \in E} m(x) < \infty$ and $\sup_{x \in E} \mathcal{V}[\varphi](x) < \infty$.

Recalling η^φ , introduced in (4.7), we also note that, for $x \in E$ and $0 \leq u \leq t$, we have

$$\eta_x^\varphi \left[\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle \right] = \mathcal{E}_x \left[\frac{\langle \varphi, \mathcal{Z} \rangle}{\mathcal{E}_x[\langle \varphi, \mathcal{Z} \rangle]} \sum_{j=1}^N \frac{\varphi(x_j)}{\langle \varphi, \mathcal{Z} \rangle} \sum_{\substack{i=1 \\ i \neq j}}^N \mathbb{E}_{\delta_{x_i}}[\langle \varphi, X_{t-u}^i \rangle] \right]$$

$$\begin{aligned}
 (6.1) \quad &= \frac{1}{\mathcal{E}_x[\langle \varphi, \mathcal{Z} \rangle]} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{i=1 \\ i \neq j}}^N \varphi(x_i) \right] \\
 &= \frac{\mathcal{V}[\varphi](x)}{\mathcal{E}_x[\langle \varphi, \mathcal{Z} \rangle]},
 \end{aligned}$$

where we have used the martingale property from (3.1) to obtain the second equality.

Next note that, thanks to the compensation formula given in [8], for example, applied to the Lévy system of the coupled process $(n_t, Y_t), t \geq 0$ with marks $(\Xi_i, i \geq 1)$ arriving at instantaneous rate

$$\rho(Y_{t-}) \eta_{Y_{t-}}^\varphi(d\Xi) dt = \gamma(Y_{t-}) \frac{m[\varphi](Y_{t-})}{\varphi(Y_{t-})} \eta_{Y_{t-}}^\varphi(d\Xi) dt,$$

(cf. (4.8) and (4.3)), we have

$$\mathbb{E}_x^\varphi \left[\sum_{i=1}^{n_t} \frac{\mathcal{V}[\varphi](Y_{T_i-})}{\mathcal{E}_{Y_{T_i-}}[\langle \varphi, \mathcal{Z} \rangle]} \right] = \mathbb{E}_x^\varphi \left[\int_0^t \frac{\gamma(Y_s) \mathcal{V}[\varphi](Y_s)}{\varphi(Y_s)} ds \right].$$

Combining this with (4.6) and (6.1), we obtain

$$(6.2) \quad \mathbb{E}_{\delta_x}^\varphi \left[\frac{\langle \varphi, X_t \rangle}{t} \right] - \mathbb{E}_{\delta_x}^\varphi \left[\frac{\varphi(Y_t)}{t} \right] = \frac{1}{t} \mathbb{E}_x^\varphi \left[\int_0^t \frac{\gamma(Y_s) \mathcal{V}[\varphi](Y_s)}{\varphi(Y_s)} ds \right],$$

To show that (1.12) holds, due to assumption (G1), we have $\bar{B} := \sup_{x \in E} B(x) \vee 1 < \infty$. Due to this, Fubini’s theorem and the change of measure (4.5), we have

$$\begin{aligned}
 \mathbb{E}_x^\varphi \left[\int_0^t \frac{\gamma(Y_s) \mathcal{V}[\varphi](Y_s)}{\varphi(Y_s)} ds \right] &= \int_0^t \frac{1}{\varphi(x)} \mathbb{E}_x \left[e^{\int_0^s B(Y_u) du} \varphi(Y_s) \frac{\gamma(Y_s) \mathcal{V}[\varphi](Y_s)}{\varphi(Y_s)} ds \right] \\
 &\leq e^{\bar{B}t} \int_0^t \frac{1}{\varphi(x)} \mathbb{E}_x [\gamma(Y_s) \mathcal{V}[\varphi](Y_s)] ds \\
 &\leq e^{\bar{B}t} \int_0^t (C e^{-\varepsilon s} + \Sigma) ds,
 \end{aligned}$$

where we have used assumption (G2) in the final line. Due to (1.11), the right-hand side above is finite and hence, combining this with (6.2), (3.3) (equivalently, (1.12)) holds for $k = 1$.

To prove (3.4) (equivalently, (1.13)), returning to the right-hand side of (6.2), changing variables so that $s = ut$ and noting that the spine term will converge to 0, we can use (6.2) to get

$$\begin{aligned}
 (6.3) \quad &\lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbb{E}_{\delta_x}^\varphi \left[\frac{\langle \varphi, X_t \rangle}{t} \right] - \Sigma \right| = \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbb{E}_x^\varphi \left[\int_0^1 \frac{\gamma(Y_{ut}) \mathcal{V}[\varphi](Y_{ut})}{\varphi(Y_{ut})} du \right] - \Sigma \right| \\
 &= \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \int_0^1 \mathbb{P}_{ut}^\varphi[g](x) du - \langle \varphi \tilde{\varphi}, g \rangle \right| \\
 &= 0,
 \end{aligned}$$

where we have used Theorem 5.1 with $F(x, u, t) = g(x) = \gamma(x) \mathcal{V}[\varphi](x) / \varphi(x)$.

Block 2: Decomposing the kth power of the sum in (4.6), assuming the theorem for $k - 1$.

We now assume that the theorem holds for all powers less than or equal to $k - 1$ and proceed to show that the result holds true for k . We can appeal to the spine decomposition (4.6) and the adaptation of the multinomial formula for expectations of i.i.d. random variables

in Lemma A.2 in the Appendix to write

$$\begin{aligned}
 & \frac{1}{t^k} \mathbb{E}_{\delta_x}^\varphi [(\langle \varphi, X_t \rangle - \varphi(Y_t))^k] \\
 (6.4) \quad & = (\Sigma/2)^k \mathbb{E}_{\delta_x}^\varphi \left[\sum_{j=1}^k 2^j \sum_{[k_1, \dots, k_j]_+} \binom{k}{k_1, \dots, k_j} \prod_{i=1}^j (k_i - 1)! \right. \\
 & \quad \left. \times \sum_{n=j}^\infty \frac{n!}{j!(n-j)!} \mathbf{1}_{\{n_t=n\}} \prod_{i=1}^j \frac{\eta_{Y_{T_i-}}^\varphi [\Xi_i(Y_{T_i-}, t - T_i)^{k_i}]}{(k_i - 1)!(\Sigma/2)^{k_i-1} \Sigma t^{k_i}} \right],
 \end{aligned}$$

where we recall the notation for η_x^φ was given in (4.7). The sum taken over the set $[k_1, \dots, k_j]_+$, which arises from the aforementioned result in the Appendix, consists of all combinations of strictly positive $\{k_1, \dots, k_j\}$ such that $\sum_{i=1}^j k_i = k$.

Next let $F_i(y, u)$, $i \geq 1$, be bounded measurable functions on the $E \times [0, \infty)$. By conditioning on filtration generated by the information along the spine at each time of a new branching event

$$\begin{aligned}
 & \mathbb{E}_{\delta_x}^\varphi \left[\mathbf{1}_{\{n_t=n\}} \prod_{i=1}^n F_i(Y_{T_i-}, T_i) \right] \\
 & = \mathbb{E}_{\delta_x}^\varphi \left[\int_0^t \mathbf{1}_{\{T_{n-1} \leq s_n, n_{s_n} = n-1\}} \rho(Y_{s_n}) e^{-\int_{T_{n-1}}^{s_n} \rho(Y_u) du} \right. \\
 (6.5) \quad & \quad \left. \times \prod_{i=1}^{n-1} F_i(Y_{T_i-}, T_i) F_n(Y_{s_n}, s_n) e^{-\int_{s_n}^t \rho(Y_u) du} ds_n \right] \\
 & = \mathbf{E}_x^\varphi \left[e^{-\int_0^t \rho(Y_u) du} \int_0^t \dots \int_0^t \mathbf{1}_{\{0 \leq s_1 \leq \dots \leq s_n\}} \prod_{i=1}^n F_i(Y_{s_i}, s_i) \rho(Y_{s_i}) ds_1 \dots ds_n \right] \\
 & = \frac{1}{n!} \mathbf{E}_x^\varphi \left[e^{-\int_0^t \rho(Y_u) du} \prod_{i=1}^n \left(\int_0^t F_i(Y_s, s) \rho(Y_s) ds \right) \right]
 \end{aligned}$$

where the second inequality follows by induction and the final inequality follows by considering all the permutations of the order of integration and the symmetry of the function $\prod_{i=1}^n F_i(\chi_i, s_i)$; see [8] and [32] for related calculations. We also note that the left-limits have been omitted as they are not ‘seen’ by Lebesgue measure. With (6.5) in hand, continuing the calculation in (6.4) and noting that we are effectively working with $F_{j+1} = \dots = F_n = 1$, we get

$$\begin{aligned}
 & \frac{1}{t^k} \mathbb{E}_{\delta_x}^\varphi [(\langle \varphi, X_t \rangle - \varphi(Y_t))^k] \\
 & = (\Sigma/2)^k \sum_{j=1}^k 2^j \sum_{[k_1, \dots, k_j]_+} \frac{1}{j!} \binom{k}{k_1, \dots, k_j} \prod_{i=1}^j (k_i - 1)! \\
 & \quad \times \mathbf{E}_{\delta_x}^\varphi \left[\sum_{n=j}^\infty \frac{e^{-\int_0^t \rho(Y_u) du}}{(n-j)!} \left(\int_0^t \rho(Y_u) du \right)^{n-j} \right. \\
 (6.6) \quad & \quad \left. \times \prod_{i=1}^j \frac{\frac{1}{t} \int_0^t \rho(Y_u) \eta_{Y_u}^\varphi [(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle)^{k_i}] du}{(k_i - 1)!(\Sigma/2)^{k_i-1} \Sigma t^{k_i-1}} \right]
 \end{aligned}$$

$$\begin{aligned}
 &= (\Sigma/2)^k \sum_{j=1}^k 2^j \sum_{[k_1, \dots, k_j]_+} \frac{1}{j!} \binom{k}{k_1, \dots, k_j} \prod_{i=1}^j (k_i - 1)! \\
 &\quad \times \mathbf{E}_x^\varphi \left[\prod_{i=1}^j \frac{\frac{1}{t} \int_0^t \rho(Y_u) \eta_{Y_u}^\varphi [(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle)^{k_i}] du}{(k_i - 1)! (\Sigma/2)^{k_i-1} \Sigma t^{k_i-1}} \right].
 \end{aligned}$$

In the next two blocks, we develop the $\eta_{Y_u}^\varphi [\dots]$ terms further and obtain asymptotics for them that allow us to complete the proof in the final block.

Block 3: Developing the $\eta_{Y_u}^\varphi [\dots]$ terms in (6.6), assuming the theorem for $k - 1$.

We need to control the terms

$$\eta_{Y_u}^\varphi \left[\left(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle \right)^{k_i} \right]$$

in (6.6), noting, in particular, that none of the k_i -s will exceed the value k .

In a similar spirit to (6.1), we have, for any integer $1 \leq p \leq k$,

$$\begin{aligned}
 (6.7) \quad &\eta_x^\varphi \left[\left(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle \right)^p \right] \\
 &= \mathcal{E}_x \left[\frac{\langle \varphi, \mathcal{Z} \rangle}{\mathcal{E}_x[\langle \varphi, \mathcal{Z} \rangle]} \sum_{j=1}^N \frac{\varphi(x_j)}{\langle \varphi, \mathcal{Z} \rangle} \bigotimes_{i=1}^N \mathbb{E}_{\delta_{x_i}} \left[\left(\sum_{\substack{i=1 \\ i \neq j}}^N \langle \varphi, X_{t-u}^i \rangle \right)^p \right] \right] \\
 &= \frac{1}{\mathcal{E}_x[\langle \varphi, \mathcal{Z} \rangle]} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{p_j=0 \\ [p_1, \dots, p_N]}} \binom{p}{p_1, \dots, p_N} \prod_{i=1}^N \mathbb{E}_{\delta_{x_i}} [\langle \varphi, X_{t-u} \rangle^{p_i}] \right],
 \end{aligned}$$

where in the final equality we have used the multinomial formula in Lemma A.1 of the Appendix; in particular, the sum is over the set $[p_1, \dots, p_N]$ of all combinations of $(p_i, i = 1, \dots, N)$ such that $\sum_{i=1}^N p_i = p$.

We start by noting that there are no more than p strictly positive elements in the product on the right-hand side above, since, regardless of the value of N , the set $\{1, \dots, p\}$ cannot be partitioned into more than p nonempty blocks. Then, the induction hypothesis (set as valid for $k - 1$), the fact that $\sum_{[p_1, \dots, p_N]} \binom{p}{p_1, \dots, p_N} = N^p$, the boundedness of φ and the assumption (1.11) ensures that, uniformly in x , the entirety of the right-hand side, and hence the left-hand side, of (6.7) is finite.

To show that this implies (1.12), we first note that, using the same techniques as in the proof of Theorem 5.1, one can show that, under the assumptions of this theorem, there exists a constant $K > 0$ such that

$$\sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{i=1}^j \int_0^1 F_i(Y_{ut}, u, t) du \right] - \prod_{i=1}^j \int_0^1 \langle \varphi \tilde{\varphi}, F_i(\cdot, u) \rangle du \right| \leq K.$$

Then, feeding (6.7) into the right-hand side of (6.6), given the definition of ρ in (4.3), and the above inequality, we see that (1.12) holds.

Block 4: Uniform asymptotics of $\eta_{Y_u}^\varphi [\dots]$ terms in (6.6) assuming the theorem for $k - 1$.

It follows from (6.7) that, for $t \geq 1$,

$$\sup_{u \leq t} \left| \frac{\mathcal{E}_x[\langle \varphi, \mathcal{Z} \rangle]}{t^{k-1}} \eta_x^\varphi \left[\left(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle \right)^k \right] \right|$$

$$\begin{aligned}
 & - \frac{1}{t^{k-1}} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{i=1 \\ i \neq j}}^N \mathbb{E}_{\delta_{x_i}} [\langle \varphi, X_{t-u} \rangle^k] \right] \\
 & \leq \frac{1}{t} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{k_j=0 \\ [k_1, \dots, k_N]_2}} \binom{k}{k_1, \dots, k_N} \prod_{i=1}^N \sup_{u \leq t} \mathbb{E}_{\delta_{x_i}} [\langle \varphi, X_s \rangle^{k_i}] / t^{k_i-1} \right] \\
 & \leq \frac{1}{t} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{k_j=0 \\ [k_1, \dots, k_N]_2}} \binom{k}{k_1, \dots, k_N} \prod_{i=1}^N \left(\frac{k_i}{k_i-1} \right)^{k_i} \varphi(x_i) \frac{\mathbb{E}_{\delta_{x_i}}^\varphi [\langle \varphi, X_t \rangle^{k_i-1}]}{t^{k_i-1}} \right],
 \end{aligned}$$

where $[k_1, \dots, k_N]_2$ are the combinations of nonnegative integers k_1, \dots, k_N such that $\sum_{i=1}^N k_i = k$ with at least two of the $k_i > 0$. Using similar arguments to the treatment of the right-hand side of (6.7), the assumption (1.11), the fact that $\|\varphi\|_\infty < \infty$ and the induction hypothesis at $k - 1$ gives us that, uniformly in $x \in E$, the right-hand side tends to zero. That is to say

$$\begin{aligned}
 (6.8) \quad & \lim_{t \rightarrow \infty} \sup_{x \in E} \sup_{u \leq t} \left| \frac{\mathcal{E}_x[\langle \varphi, \mathcal{Z} \rangle]}{t^{k-1}} \eta_x^\varphi \left[\left(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle \right)^k \right] \right. \\
 & \left. - \frac{1}{t^{k-1}} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{i=1 \\ i \neq j}}^N \mathbb{E}_{\delta_{x_i}} [\langle \varphi, X_{t-u} \rangle^k] \right] \right| = 0.
 \end{aligned}$$

This gives the uniform leading order behaviour of the term on the left-hand side of (6.7).

Block 5: Uniform asymptotic martingale moments assuming the theorem for $k - 1$.

If we can show that, for each $j \leq k$,

$$(6.9) \quad \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{\ell=1}^j \frac{\frac{1}{t} \int_0^t du \rho(Y_u) \eta_{Y_u}^\varphi [(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle)^{k_\ell}]}{(k_\ell - 1)! (\Sigma/2)^{k_\ell-1} \Sigma t^{k_\ell-1}} \right] - 1 \right| = 0,$$

then this, Lemma A.3 in the Appendix and (6.6) yield

$$\begin{aligned}
 \lim_{t \rightarrow \infty} \frac{1}{t^k} \mathbb{E}_{\delta_x}^\varphi \left[\left(\langle \varphi, X_t \rangle - \varphi(Y_t) \right)^k \right] &= \frac{1}{t^k} \mathbb{E}_{\delta_x}^\varphi [\langle \varphi, X_t \rangle^k] \\
 &= (\Sigma/2)^k \sum_{j=1}^k 2^j \sum_{[k_1, \dots, k_j]_+} \frac{1}{j!} \binom{k}{k_1, \dots, k_j} \prod_{i=1}^j (k_i - 1)! \\
 &= (\Sigma/2)^k (k + 1)!
 \end{aligned}$$

uniformly for $x \in E$ which completes the proof of the theorem. Note that we have used the fact that $\varphi(Y_t)/t \rightarrow 0$ as $t \rightarrow \infty$ in the first equality above.

The remainder of the proof is thus focused on showing (6.9). Suppose that $1 \leq j \leq k$ and $k_1, \dots, k_j \in \mathbb{N}$ are such that $\sum_{i=1}^j k_i = k$. Then, from (6.8), used in conjunction with (6.6), we have

$$\lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{\ell=1}^j \frac{\frac{1}{t} \int_0^t du \rho(Y_u) \eta_{Y_u}^\varphi [(\sum_{\substack{i=1 \\ i \neq i^*}}^N \langle \varphi, X_{t-u}^i \rangle)^{k_\ell}]}{(k_\ell - 1)! (\Sigma/2)^{k_\ell-1} \Sigma t^{k_\ell-1}} \right] - 1 \right|$$

$$(6.10) = \lim_{t \rightarrow \infty} \sup_{x \in E} \left| \mathbf{E}_x^\varphi \left[\prod_{\ell=1}^j \frac{\frac{1}{t} \int_0^t du \frac{\gamma(Y_u)}{\varphi(Y_u)} \mathcal{E}_{Y_u} [\sum_{j=1}^N \varphi(Y_j) \sum_{\substack{i=1 \\ i \neq j}}^N \mathbb{E}_{\delta_{Y_i}} [\langle \varphi, X_{t-u} \rangle^{k_\ell}]]]}{(k_\ell - 1)! (\Sigma/2)^{k_\ell - 1} \Sigma t^{k_\ell - 1}} \right] - 1 \right|$$

$$= \lim_{t \rightarrow \infty} \sup_{x \in E} \frac{1}{\Sigma^j} \left| \mathbf{E}_x^\varphi \left[\prod_{\ell=1}^j \int_0^1 ds F_\ell(Y_{st}, s, t) \right] - \Sigma^j \right|,$$

where, for $\ell = 1, \dots, j, x \in E, 0 \leq u \leq 1$ and $t \geq 0$,

$$F_\ell(x, s, t) = \frac{\gamma(x)}{\varphi(x)} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{i=1 \\ i \neq j}}^N \varphi(x_i) \frac{\mathbb{E}_{\delta_{x_i}}^\varphi [\langle \varphi, X_{t(1-s)} \rangle^{k_\ell - 1}]}{(k_\ell - 1)! (\Sigma/2)^{k_\ell - 1} t^{k_\ell - 1}} \right]$$

$$= k_\ell (1-s)^{k_\ell - 1} \frac{\gamma(x)}{\varphi(x)} \mathcal{E}_x \left[\sum_{j=1}^N \varphi(x_j) \sum_{\substack{i=1 \\ i \neq j}}^N \varphi(x_i) \frac{\mathbb{E}_{\delta_{x_i}}^\varphi [\langle \varphi, X_{t(1-s)} \rangle^{k_\ell - 1}]}{k_\ell! (\Sigma/2)^{k_\ell - 1} (t(1-s))^{k_\ell - 1}} \right].$$

Since we are assuming that the statement of the theorem holds for $k - 1$, it is easy to verify with the help of (1.11) that $F_i \in \mathcal{C}$ (where \mathcal{C} was defined immediately before Theorem 5.1) and that

$$F_\ell(x, s) := \lim_{t \rightarrow \infty} F_\ell(x, s, t) = k_\ell (1-s)^{k_\ell - 1} \frac{\gamma(x)}{\varphi(x)} \gamma[\varphi](x).$$

Note also that $k_\ell (1-s)^{k_\ell - 1} \Sigma = \langle \varphi \tilde{\varphi}, F_\ell(\cdot, s) \rangle$ so that $\int_0^1 \langle \varphi \tilde{\varphi}, F_\ell(\cdot, s) \rangle ds = \Sigma$. As a consequence, Theorem 5.1 tells us that the right-hand side of (6.10) converges to zero, as required. □

7. Step 2: Proof of Theorem 1.2. In addition to (G1) and (G2), assumed in the previous section, we further assume (G3), (G4) and (G5). Note that (G5) implies that the previously enforced moment condition (1.11) holds.

For bounded functions $g : E \rightarrow [0, 1]$, let us define the nonlinear semigroup,

$$u_t[g](x) = \mathbb{E}_{\delta_x} \left[1 - \prod_{i=1}^{N_t} g(x_i(t)) \right], \quad x \in E,$$

with the convention that if $\zeta \leq t$, then the above product is equal to 1. We will also use the notation

$$(7.1) \quad u_t(x) := u_t[\mathbf{0}](x) = \mathbb{P}_{\delta_x}(\zeta > t) \quad \text{and} \quad a(t) := \langle \tilde{\varphi}, u_t \rangle,$$

where $\mathbf{0}$ is the zero function.

PROPOSITION 7.1. *For all bounded and measurable functions $g \in L_\infty^+(E)$ such that $\|g\|_\infty \leq 1$ and $x \in E, t \geq 0, u_t[g]$ satisfies*

$$(7.2) \quad u_t[g](x) = \psi_t[1 - g](x) + \int_0^t \psi_s[\mathcal{G}[u_{t-s}[g]]](x) ds,$$

where the operator \mathcal{G} is given by

$$(7.3) \quad \mathcal{G}[h](x) := \gamma(x) \mathcal{E}_x \left[1 - \prod_{i=1}^N (1 - h(x_i)) - \sum_{i=1}^N h(x_i) \right]$$

for $h \in L_{\infty,1}^+(E)$. It follows that $a(t)$, defined in (7.1), satisfies

$$(7.4) \quad a(t) = a(0) + \int_0^t \langle \tilde{\varphi}, \mathcal{G}[u_{t-s}] \rangle ds = a(0) + \int_0^t \langle \tilde{\varphi}, \mathcal{G}[u_s] \rangle ds, \quad t \geq 0.$$

PROOF. Again, the proof uses the usual method of splitting on the first branching event; see, for example, [11, 20, 26], so we will only sketch the arguments here. Recall that $(P_t, t \geq 0)$ is the underlying motion and that the branching operator \mathcal{A} given in (2.2). By conditioning on the first branching event and then applying Lemma 1.2, Chapter 4 in [17], we find that, for $x \in E, t \geq 0$ and $g \in L_{\infty,1}^+(E)$,

$$(7.5) \quad u_t[g](x) = P_t[1 - g](x) + \int_0^t P_s[\mathcal{A}[u_{t-s}[g]]](x) ds + \int_0^t P_s[\mathcal{G}[u_{t-s}[g]]](x) ds.$$

Next, recall the process $(Y_t, t \geq 0)$ under \mathbf{P}_x defined just before Lemma 2.1. Define $(\tilde{u}_t, t \geq 0)$ via

$$(7.6) \quad \tilde{u}_t[g](x) = \mathbf{E}_x[e^{\int_0^t B(Y_s) ds}(1 - g(Y_t))] + \mathbf{E}_x\left[\int_0^t e^{\int_0^s B(Y_u) du} \mathcal{G}[\tilde{u}_{t-s}[g]](Y_s)\right],$$

for $x \in E, t \geq 0$ and $g \in L_{\infty,1}^+(E)$. Note that for the moment we don't claim a solution to (7.6) exists. By conditioning the right-hand side of (7.6) on the first jump of Y and invoking Lemma 1.2, Chapter 4 of [17] (bearing in mind the dynamics of Y given just before Lemma 2.1), we can check that $(\tilde{u}_t, t \geq 0)$ is a solution to (7.6) if and only if it solves (7.5). In particular, $(u_t, t \geq 0)$ satisfies (7.6). Combined with the many-to-one formula, this implies that

$$(7.7) \quad u_t[g](x) = \psi_t[1 - g](x) + \int_0^t \psi_s[\mathcal{G}[u_{t-s}[g]]](x) ds$$

which is (7.2). As $\tilde{\varphi}$ is a left eigenfunction for the semigroup $(\psi_t, t \geq 0)$ and as we are dealing with the critical case, we have, for all $t \geq 0$, that $\langle \tilde{\varphi}, \psi_t[g] \rangle = \langle \tilde{\varphi}, g \rangle$. Using this fact and taking $g = \mathbf{1}$ in (7.7), we find that

$$a(t) = \langle \tilde{\varphi}, \psi_t[\mathbf{1}] \rangle + \int_0^t \langle \tilde{\varphi}, \psi_s[\mathcal{G}[u_{t-s}]] \rangle ds = \langle \tilde{\varphi}, \mathbf{1} \rangle + \int_0^t \langle \tilde{\varphi}, \mathcal{G}[u_{t-s}] \rangle ds,$$

where $\mathbf{1}$ is the function which is identically unity. The above equality gives rise to the first identity in (7.4). The second identity in (7.4) then follows by a change of variable. \square

Our aim is to extract the proof of Theorem 1.2 from (7.2) and (7.4). To this end, we need a series of technical lemmas. The first lemma regards the analytical behaviour of the $\mathcal{V}[h](x)$, defined in (1.9), and $\mathcal{G}[h](x)$ defined in (7.3).

LEMMA 7.1. *The following statements hold:*

(i) *For all $x \in E$ and $h : E \rightarrow [0, 1]$, we have $0 \geq \mathcal{G}[h](x) \geq -\|\gamma\|_{\infty} M$, where $M = \sup_{x \in E} m(x)$.*

(ii) *There exists $C \in (0, \infty)$ such that, for all $h : E \rightarrow [0, \frac{1}{2}]$,*

$$(7.8) \quad \sup_{x \in E} \left| \mathcal{G}[h](x) + \frac{1}{2} \gamma(x) \mathcal{V}[h](x) \right| \leq C \sup_{x \in E} h(x)^3.$$

(iii) *There exists $C \in (0, \infty)$ such that, for all $h : E \rightarrow [0, \frac{1}{2}]$,*

$$-\mathcal{G}[h](x) \geq C \gamma(x) \mathcal{V}[h](x) \geq 0, \quad x \in E.$$

(iv) *There exists $C \in (0, \infty)$ such that, for all $h_1, h_2 : E \rightarrow [0, 1]$, we have*

$$\sup_{x \in E} |\mathcal{V}[h_1](x) - \mathcal{V}[h_2](x)| \leq C \sup_{x \in E} |h_1(x) - h_2(x)|.$$

PROOF. (i) The nonpositivity of $\mathcal{G}[h]$ can be demonstrated using an inductive argument. For $n \geq 1$ and $(x_i, 1 \leq i \leq n) \in [0, 1]^n$,

$$f_n(x_1, \dots, x_n) := 1 - \prod_{i=1}^n (1 - x_i) - \sum_{i=1}^n x_i \leq 0.$$

To see why this is true, we note that $f_1(x_1) = 0$, and more generally, we have $f_{n+1}(x_1, \dots, x_n, 0) = f_n(x_1, \dots, x_n)$ as well as $\partial_{n+1} f_{n+1}(x_1, \dots, x_{n+1}) \leq 0$. This yields $\mathcal{G}[h] \leq 0$.

For the second inequality in (i), it suffices to observe that, since $0 \leq h \leq 1$,

$$\mathcal{E}_x \left[1 - \prod_i (1 - h(x_i)) - \langle h, \mathcal{Z} \rangle \right] \geq -\mathcal{E}_x[\langle h, \mathcal{Z} \rangle] \geq -\sup_{x \in E} m(x),$$

which is bounded due to (G1).

(ii) Let us write $L(h) = -\langle \log(1 - h), \mathcal{Z} \rangle$ which is nonnegative. Then, we have

$$\begin{aligned} (7.9) \quad \left| \mathcal{G}[h](x) + \frac{1}{2} \gamma(x) \mathcal{V}[h](x) \right| &\leq \gamma(x) \mathcal{E}_x \left[1 - e^{-L(h)} - L(h) + \frac{1}{2} L(h)^2 \right] \\ &+ \gamma(x) \mathcal{E}_x \left[\left\langle -\log(1 - h) - h - \frac{1}{2} h^2, \mathcal{Z} \right\rangle \right] \\ &+ \frac{1}{2} \gamma(x) \mathcal{E}_x [|\langle \log(1 - h), \mathcal{Z} \rangle|^2 - \langle h, \mathcal{Z} \rangle^2]. \end{aligned}$$

Applying the elementary bounds that $0 \leq 1 - e^{-x} - x + \frac{1}{2}x^2 \leq \frac{1}{6}x^3$ for all $x \geq 0$ and that $0 \leq -\log(1 - x) - x - \frac{1}{2}x^2 \leq x^3$ for $x \in [0, \frac{1}{2}]$, since both γ and the number of offspring produced at a branching event are uniformly bounded, thanks to (G1) and (G4), respectively, we see the first two terms on the right-hand side of (7.9) are bounded by $C \|h\|_\infty^3$. For the third term on the right-hand side of (7.9), writing $\log(1 - h) = -h + R(h)$ with $|R(h)| \leq C \|h\|_\infty^2$, we get

$$|\langle \log(1 - h), \mathcal{Z} \rangle^2 - \langle h, \mathcal{Z} \rangle^2| \leq 2 \langle h, \mathcal{Z} \rangle \cdot \langle |R(h)|, \mathcal{Z} \rangle + \langle R(h), \mathcal{Z} \rangle^2 \leq C \|h\|_\infty^3,$$

which leads to (ii).

(iii) First note that, for all $n \geq 2$ and $(x_i, 1 \leq i \leq n) \in [0, \frac{1}{2}]^n$,

$$\prod_{i=1}^n (1 - x_i) - 1 + \sum_{i=1}^n x_i \geq \frac{1}{2^{n-1}} \sum_{1 \leq i, j \leq n: i \neq j} x_i x_j.$$

This can be shown with an induction on n and its proof is, therefore, omitted. The result then follows due to (G4).

(iv) The final claim can be checked from the definition of \mathcal{V} , combined with (G4). \square

We now prove a sequence of results that show that $a(t)$ and $u_t(x)$ can be bounded above and below by a constant divided by t . We can then use this to show that the difference between $u_t(x)/\varphi$ and $a(t)$ decays sufficiently quickly as $t \rightarrow \infty$, which we then, in turn, use to obtain the precise constant given in Theorem 1.2. The proof is inspired by Section 5 of [33], where Theorem 1.2 was proved for the NBP in the isotropic case with homogeneous branching. In this case, equation (7.1) and the operator \mathcal{V} simplify due to the lack of correlation between particles produced at branching events. This is not possible in the general case, as particles may be correlated.

We start with the lower bound.

LEMMA 7.2. *There exists $C \in (0, \infty)$ such that*

$$\mathbb{P}_{\delta_x}(\zeta > t) \geq \frac{\varphi(x)}{Ct} \quad \text{and} \quad a(t) \geq \frac{1}{Ct}$$

for all $t \geq 1$.

PROOF. Let us recall the change of measure (4.1) and that, from Theorem 1.1, there exists a $C > 0$ such that

$$\sup_{x \in E} \mathbb{E}_{\delta_x}^\varphi [\langle \varphi, X_t \rangle] \leq Ct$$

for all $t \geq 1$. By Jensen’s inequality we then get

$$(7.10) \quad \mathbb{P}_{\delta_x}(\zeta > t) = \mathbb{E}_{\delta_x}^\varphi \left[\frac{\varphi(x)}{\langle \varphi, X_t \rangle} \right] \geq \frac{\varphi(x)}{\mathbb{E}_{\delta_x}^\varphi [\langle \varphi, X_t \rangle]} \geq \frac{\varphi(x)}{Ct}, \quad t \geq 1.$$

The lower bound for $a(t)$ then follows from an integration with $\tilde{\varphi}$, recalling that we have normalised the left and right eigenfunctions so that $\langle \tilde{\varphi}, \varphi \rangle = 1$. \square

Before turning to the upper bound, we first state and prove a lemma regarding the extinction time of the process.

LEMMA 7.3. *Under the assumptions of Theorem 1.2, there exists a constant $C > 0$ such that, for all $t \geq t_0$,*

$$(7.11) \quad a(t) \leq \frac{C}{t} \quad \text{and} \quad \sup_{x \in E} u_t(x) \leq \frac{C}{t}.$$

PROOF. We first show that

$$a(t) \rightarrow 0 \quad \text{and} \quad \sup_{x \in E} u_t(x) \rightarrow 0$$

as $t \rightarrow \infty$.

Recall $\zeta = \inf\{t > 0 : \langle 1, X_t \rangle = 0\}$ is the extinction time. Due to (G3), we have $\mathbb{P}_{\delta_x}(\zeta < \infty) = 1$, namely, $u_t(x) = \mathbb{P}_{\delta_x}(\zeta > t) \rightarrow 0$, as $t \rightarrow \infty$, for all $x \in E$. Therefore, $a(t) = \langle \tilde{\varphi}, u_t \rangle \rightarrow 0$, as $t \rightarrow \infty$, by dominated convergence. To get to the uniform convergence of u_t to zero, let us note that $u_{t+s}(x) = u_t[1 - u_s](x)$ by the Markov branching property. Appealing to (7.2) and Lemma 7.1(i), we find that

$$(7.12) \quad 0 \leq u_{t+s}(x) = \psi_t[u_s](x) + \int_0^t \psi_l[\mathcal{G}[u_{t+s-l}]](x) dl \leq \psi_t[u_s](x)$$

by Lemma 7.1(i). Combined with (G2), this yields

$$(7.13) \quad \|u_{t+s}\|_\infty \leq \|\psi_t[u_s]\|_\infty \leq a(s)\|\varphi\|_\infty + O(e^{-\varepsilon t}).$$

Taking first $t \rightarrow \infty$ and then $s \rightarrow \infty$ gives the desired limit $\|u_t\| \rightarrow 0$ as $t \rightarrow 0$.

Now, we prove the required upper bound on $a(t)$ and $\|u_t\|_\infty$. Due to the above calculations, we may fix $t_0 > 0$ such that $\|u_t\|_\infty \leq 1/2$ for all $t \geq t_0$. Note that the integrand in (7.4) is bounded due to our assumptions and Lemma 7.1(i). It follows that $a(t)$ is differentiable. Differentiating (7.4) for $t \geq t_0$ and then applying the bound in Lemma 7.1(iii), we obtain that, for $t \geq t_0$,

$$a'(t) = \langle \tilde{\varphi}, \mathcal{G}[u_t] \rangle \leq -C_1 \langle \tilde{\varphi}, \gamma \mathcal{V}[u_t] \rangle \leq -C_2 \langle \tilde{\varphi}, u_t \rangle^2 = -C_2 a(t)^2,$$

where we have used assumption (G5) in the second inequality.

Integrating from t_0 to t yields

$$a(t) \leq (C_2(t - t_0) + a(t_0)^{-1})^{-1} \leq (C_2t)^{-1},$$

where the last inequality holds for t sufficiently large. The upper bound for $a(t)$ then follows. We may then apply the same techniques as in (7.13) by setting $s = t$ to obtain the uniform bound for $u_t(x)$. \square

The next result shows that the long-term behaviour of u_t/φ and $a(t)$ are the same which will be key to obtaining the correct constants in the bounds obtained in the previous lemma.

LEMMA 7.4. *Suppose that there exist $\kappa, \eta \in (0, \infty)$ such that $\|u_t\|_\infty \leq \kappa t^{-\eta}$ for all $t > 0$. Then, we can find some constant $C \in (0, \infty)$ which does not depend on κ such that*

$$\sup_{x \in E} \left| \frac{u_t(x)}{\varphi(x)} - a(t) \right| \leq C\kappa^2 t^{-2\eta} \quad \text{for all } t > 0.$$

PROOF. Comparing (7.2) with (7.4), we find that

$$\begin{aligned} \left| \frac{u_t(x)}{\varphi(x)} - a(t) \right| &\leq \left| \frac{\psi_t[\mathbf{1}](x)}{\varphi(x)} - \langle \tilde{\varphi}, \mathbf{1} \rangle \right| + \int_0^t \left| \frac{\psi_{t-s}[\mathcal{G}[u_s]](x)}{\varphi(x)} - \langle \tilde{\varphi}, \mathcal{G}[u_s] \rangle \right| ds \\ &\leq C_1 e^{-\varepsilon t} + \int_0^t C_1 e^{-\varepsilon(t-s)} \|\mathcal{G}[u_s]\|_\infty ds, \end{aligned}$$

where the constant $\varepsilon > 0$ exists thanks to (G2). Since we assumed that $\|u_t\|_\infty \leq \kappa t^{-\eta}$ for some constants κ, η , we can find $t_0 > 0$ such that $\sup_{x \in E} \sup_{t \geq t_0} u_t(x) \leq 1/2$. Take $t \geq 2t_0$, then the integral above can be bounded as follows:

$$\begin{aligned} &\int_0^t C_1 e^{-\varepsilon(t-s)} \|\mathcal{G}[u_s]\|_\infty ds \\ (7.14) \quad &= \int_0^{t/2} C_1 e^{-\varepsilon(t-s)} \|\mathcal{G}[u_s]\|_\infty ds + \int_{t/2}^t C_1 e^{-\varepsilon(t-s)} \|\mathcal{G}[u_s]\|_\infty ds \\ &\leq \frac{C_1 \|\mathcal{V}\|_\infty N_{\max}}{\varepsilon} e^{-\varepsilon t/2} + \int_{t/2}^t C_1 e^{-\varepsilon(t-s)} \|\mathcal{G}[u_s]\|_\infty ds, \end{aligned}$$

where we have used Lemma 7.1(i). We note that, due to (G4), $\mathcal{V}[u_s] \leq N_{\max}^2 \|u_s\|_\infty^2$. Using this together with Lemma 7.1(ii), we deduce for the second term in (7.14) that $\|\mathcal{G}[u_s]\|_\infty \leq \|\mathcal{V}[u_s]\|_\infty + \|u_s^3\|_\infty \leq C_2 \|u_s\|_\infty^2$. The latter is bounded by $C_2 \kappa^2 s^{-2\eta}$ due to the assumption of the lemma. Therefore,

$$\begin{aligned} \int_{t/2}^t C_1 e^{-\varepsilon(t-s)} \|\mathcal{G}[u_s]\|_\infty ds &\leq C_3 \kappa^2 \int_{t/2}^t e^{-\varepsilon(t-s)} s^{-2\eta} ds \\ &= \frac{C_3 \kappa^2}{\varepsilon} (t^{-2\eta} - 4e^{-\varepsilon t/2} t^{-2\eta}) + C_4 \int_{t/2}^t e^{-\varepsilon(t-s)} s^{-2\eta-1} ds \\ &= O(t^{-2\eta}). \end{aligned}$$

Putting the pieces together, we obtain the claimed bound in the lemma. \square

We are now ready to prove our Theorem 1.2.

PROOF OF THEOREM 1.2. Applying Lemma 7.4 with $\eta = 1$ and κ being some positive constant (this is permitted thanks to Lemma 7.3), we have

$$\sup_{x \in E} \left| \frac{u_t(x)}{\varphi(x)} - a(t) \right| = O(t^{-2}), \quad t \rightarrow \infty.$$

On the other hand, we've seen in Lemma 7.2 that $a(t)^{-1} = O(t)$. It follows that

$$(7.15) \quad \sup_{x \in E} \left| \frac{u_t(x)}{\varphi(x)a(t)} - 1 \right| = O(t^{-1}), \quad t \rightarrow \infty.$$

Applying Lemma 7.1(iv), we deduce that

$$(7.16) \quad \begin{aligned} & \sup_{x \in E} |\mathcal{V}[u_t](x) - a(t)^2 \mathcal{V}[\varphi](x)| \\ &= \sup_{r \in D, v \in V} a(t)^2 \left| \mathcal{V} \left[\frac{u_t}{a(t)} \right](x) - \mathcal{V}[\varphi](x) \right| \\ &\leq C a(t)^2 \sup_{x \in E} \left| \frac{u_t(x)}{a(t)} - \varphi(x) \right| = O(t^{-3}). \end{aligned}$$

Therefore, for all $t \geq t_0$,

$$\begin{aligned} a(t) - a(t_0) &= \int_{t_0}^t \langle \tilde{\varphi}, \mathcal{G}[u_s] \rangle ds \\ &= -\frac{1}{2} \int_{t_0}^t (\langle \tilde{\varphi}, \gamma \mathcal{V}[u_s] \rangle + O(\|u_s\|_\infty^3)) ds \\ &= -\frac{1}{2} \int_{t_0}^t (\langle \tilde{\varphi}, \gamma \mathcal{V}[u_s] \rangle + O(s^{-3})) ds \\ &= -\frac{1}{2} \int_{t_0}^t (a(s)^2 \langle \tilde{\varphi}, \gamma \mathcal{V}[\varphi] \rangle + O(s^{-3})) ds \\ &= -\frac{\Sigma}{2} \int_{t_0}^t a(s)^2 (1 + o(1)) ds, \end{aligned}$$

where we have used (7.8) in the second equality, (7.11) in the third, (7.16) in the fourth and Lemma 7.2 in the final equality. This implies that

$$a(t) \sim \frac{2}{\Sigma t} \quad \text{as } t \rightarrow \infty.$$

The desired asymptotic for u_t then follows from (7.15). \square

8. Step 3: Proof of Theorem 1.3. Returning to (3.5), the conclusions of Theorem 1.1 and Theorem 1.2 yield

$$\lim_{t \rightarrow \infty} \mathbb{E}_{\delta_x} [\langle \varphi, X_t \rangle^k | \zeta > t] = \lim_{t \rightarrow \infty} \frac{\mathbb{E}_{\delta_x}^\varphi [\langle \varphi, X_t \rangle^{k-1} / t^{k-1}]}{t \mathbb{E}_{\delta_x}^\varphi [1 / \langle \varphi, X_t \rangle]} = (\Sigma/2)^k k!, \quad k \in \mathbb{N}.$$

The right-hand side above is precisely the k th moment of an exponential distribution with parameter $2/\Sigma$. Hence, at this stage, again thanks to the *moment problem* for the exponential distribution, noting that $\langle \tilde{\varphi}, \varphi \rangle = 1$, we conclude that Theorem 1.3 holds for the special case that $f = \varphi$.

We now deal more generally with the case that $f \in L_\infty^+(E)$.

PROOF OF THEOREM 1.3. As alluded to above, we introduce $\tilde{f} = f - \langle \tilde{\varphi}, f \rangle \varphi$ and consider the behaviour of $\langle \tilde{f}, X_t \rangle / t$ as $t \rightarrow \infty$. The first thing we should note in this direction is that

$$\langle \tilde{\varphi}, \tilde{f} \rangle = \langle \tilde{\varphi}, f \rangle - \langle \tilde{\varphi}, f \rangle \langle \tilde{\varphi}, \varphi \rangle = 0,$$

where we have used that $\langle \tilde{\varphi}, \varphi \rangle = 1$. Hence, assumption (G2) tells us that, for t sufficiently large,

$$\sup_{x \in E} \left| \frac{\psi_t[\tilde{f}](x)}{\varphi(x)} \right| = O(e^{-\epsilon t})$$

for some $\epsilon > 0$. In particular, we have $\sup_{x \in E} \psi_t[\tilde{f}](x) \rightarrow 0$ as $t \rightarrow \infty$.

We also note that, under the assumption that $0 \leq f \leq \varphi$, we have the estimate

$$(8.1) \quad -\langle \tilde{\varphi}, f \rangle \varphi \leq \tilde{f} \leq (1 - \langle \tilde{\varphi}, f \rangle) \varphi.$$

Further, since $f \leq \varphi$, $\langle \tilde{\varphi}, f \rangle \leq \langle \tilde{\varphi}, \varphi \rangle = 1$, the upper bound in (8.1) is positive. As such, $|\tilde{f}| \in L^+_\infty(E)$.

We will use these facts to prove that $\langle \tilde{f}, X_t \rangle / t$ converges weakly under $\mathbb{P}_{\delta_x}(\cdot | \zeta > t)$ to zero as $t \rightarrow \infty$. In that case, thanks again to Slutsky’s Theorem, we can deduce that, for $x \in E$,

$$(8.2) \quad \begin{aligned} & \lim_{t \rightarrow \infty} \mathbb{E}_{\delta_x} \left[\exp \left(-\theta \frac{\langle f, X_t \rangle}{t} \right) \middle| \zeta > t \right] \\ &= \lim_{t \rightarrow \infty} \mathbb{E}_{\delta_x} \left[\exp \left(-\theta \frac{\langle \tilde{f}, X_t \rangle}{t} - \theta \langle \tilde{\varphi}, f \rangle \frac{\langle \varphi, X_t \rangle}{t} \right) \middle| \zeta > t \right] \\ &= \frac{1}{1 + \Sigma \theta \langle \tilde{\varphi}, f \rangle / 2}, \end{aligned}$$

as required.

Let us thus consider the estimate

$$(8.3) \quad \mathbb{P}_{\delta_x} (|\langle \tilde{f}, X_t \rangle| / t > \epsilon | \zeta > t) \leq \frac{1}{\epsilon^2} \frac{\mathbb{E}_{\delta_x} [\langle \tilde{f}, X_t \rangle^2 / t]}{t \mathbb{P}_{\delta_x} (\langle 1, X_t \rangle > 0)}.$$

In order to deal with the numerator on the right-hand side above, we will use the *many-to-two* formula given in Lemma 2.2. We restate it here in the required form for convenience.

Let $g \in L^+_\infty(E)$. Then, the many-to-two formula reads

$$(8.4) \quad \mathbb{E}_{\delta_x} [\langle g, X_t \rangle^2] = \psi_t[g^2](x) + \int_0^t \psi_s[\gamma \mathcal{V}[\psi_{t-s}[g]]](x) ds$$

for $x \in E$ and $t \geq 0$, where \mathcal{V} was given in (1.9).

Hence, setting $g = \tilde{f}$ and dividing by t^2 , we note that the first term on the right-hand side of (8.4) has the behaviour

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sup_{x \in E} |\psi_t[\tilde{f}^2](r, \nu)| \leq \lim_{t \rightarrow \infty} \frac{1}{t} \sup_{x \in E} |\psi_t[\tilde{f}^2](x) - \langle \tilde{\varphi}, \tilde{f}^2 \rangle \varphi| + \frac{1}{t} \langle \tilde{\varphi}, \tilde{f}^2 \rangle \|\varphi\|_\infty = 0,$$

where the final equality uses (G2). For the integral term in (8.4), we can handle it using the ergodic convolution Theorem 5.1, the change of measure Lemma 4.1 and the many-to-one formula in Lemma 2.1. In particular, for $u \in [0, 1]$ and $t \geq 0$,

$$\begin{aligned} \psi_{ut}[\gamma \mathcal{V}[\psi_{t(1-u)}[\tilde{f}]]](x) &= \frac{\psi_{ut}[\varphi \frac{\gamma}{\varphi} \mathcal{V}[\psi_{t(1-u)}[\tilde{f}]]](x)}{\varphi(x)} \\ &= \mathbf{E}_x^\varphi \left[\frac{\gamma(Y_{ut})}{\varphi(Y_{ut})} \mathcal{V}[\psi_{t(1-u)}[\tilde{f}]](Y_{ut}) \right] \varphi(x). \end{aligned}$$

Moreover, defining $F(x, u, t) = \varphi(x)^{-1} \gamma(r, \nu) \mathcal{V}[\psi_{t(1-u)}[\tilde{f}]](x)$, Lemma 7.1 and the fact that $\psi_t[\tilde{f}] \rightarrow 0$ imply that

$$F(x, u) := \lim_{t \rightarrow \infty} F(x, u, t) = \lim_{t \rightarrow \infty} \frac{\gamma(x)}{\varphi(x)} \mathcal{V}[\psi_{t(1-u)}[\tilde{f}]](x) = 0$$

so that the conditions of Theorem 5.1 are trivially met (i.e., $F \in \mathcal{C}$). Hence, we have

$$\lim_{t \rightarrow \infty} \mathbb{E}_{\delta_x} \left[\frac{\langle \tilde{f}, X_t \rangle^2}{t} \right] = \lim_{t \rightarrow \infty} \int_0^1 \mathbf{E}_x^\varphi \left[\frac{\gamma(Y_{ut})}{\varphi(Y_{ut})} \mathcal{V}[\psi_{t(1-u)}[\tilde{f}]](Y_{ut}) \right] du = 0.$$

Using this back in (8.3), together with the convergence of $t\mathbb{P}_{\delta_x}(\zeta > t)$ as $t \rightarrow \infty$, we see that

$$\lim_{t \rightarrow \infty} \mathbb{P}_{\delta_x}(|\langle \tilde{f}, X_t \rangle|/t > \varepsilon | \zeta > t) \leq \lim_{t \rightarrow \infty} \frac{1}{\varepsilon^2} \frac{\mathbb{E}_{\delta_x}[\langle \tilde{f}, X_t \rangle^2/t]}{t\mathbb{P}_{\delta_x}(\zeta > t)} = 0.$$

This tells us that $\langle \tilde{f}, X_t \rangle/t$ converges weakly to zero as $t \rightarrow \infty$ and hence, thanks to (8.2), this completes the proof of the Yaglom limit for general $f \in L_\infty^+(E)$. \square

9. Neutron transport equation. We will now illustrate our main results in the context of the neutron branching process (NBP) for which the nonlocal branching assumption is crucial. Moreover, we are interested in this class of branching process given the recent surge of results; see [10–12, 20, 26]. Let us spend some time discussing the underlying model. We will often reuse previously defined notation but in a way that will cause no confusion.

The neutron transport equation (NTE) describes the density of neutrons in an inhomogeneous fissile medium. The backward form of the NTE can be written as follows:¹

$$\begin{aligned} (9.1) \quad & \frac{\partial}{\partial t} \psi_t(r, \nu) \\ &= \nu \cdot \nabla_r \psi_t(r, \nu) - \sigma(r, \nu) \psi_t(r, \nu) \\ & \quad + \sigma_s(r, \nu) \int_V \psi_t(r, \nu') \pi_s(r, \nu, \nu') d\nu' + \sigma_f(r, \nu) \int_V \psi_t(r, \nu') \pi_f(r, \nu, \nu') d\nu', \end{aligned}$$

where the neutron density $\psi_t(r, \nu)$ is a function of time, $t \geq 0$, and the configuration variables $(r, \nu) \in D \times V$ where the spatial domain $D \subseteq \mathbb{R}^3$ is nonempty, smooth and bounded such that ∂D has zero Lebesgue measure, and $V = \{\nu \in \mathbb{R}^3 : \nu_{\min} \leq |\nu| \leq \nu_{\max}\}$ is the velocity space, where $0 < \nu_{\min} < \nu_{\max} < \infty$. Furthermore, the other components of (9.1) have the following interpretation:

- $\sigma_s(r, \nu)$: the rate at which scattering occurs from incoming velocity ν ,
- $\sigma_f(r, \nu)$: the rate at which fission occurs from incoming velocity ν ,
- $\sigma(r, \nu)$: the sum of the rates $\sigma_f + \sigma_s$ and is known as the total cross section,
- $\pi_s(r, \nu, \nu') d\nu'$: the scattering yield at velocity ν' from incoming velocity ν ,
satisfying $\int_V \pi_s(r, \nu, \nu') d\nu' = 1$ and
- $\pi_f(r, \nu, \nu') d\nu'$: the neutron yield at velocity ν' from fission with incoming velocity ν ,
satisfying $\int_V \pi_f(r, \nu, \nu') d\nu' < \infty$.

We also enforce the following initial and boundary conditions:

$$(9.2) \quad \begin{cases} \psi_0(r, \nu) = g(r, \nu) & \text{for } r \in D, \nu \in V, \\ \psi_t(r, \nu) = 0 & \text{for } t \geq 0 \text{ and } r \in \partial D \text{ if } \nu \cdot \mathbf{n}_r > 0, \end{cases}$$

¹The operator ∇_r corresponds to gradient with respect to the variable $r \in D$.

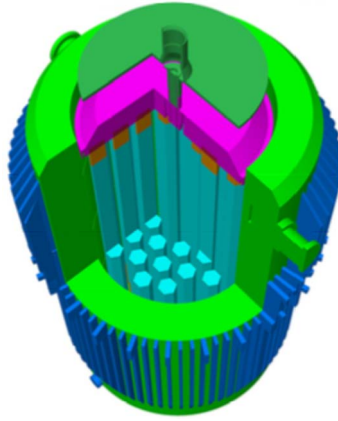


FIG. 1. The geometry of a nuclear reactor core, representing a physical domain D , onto which the different cross-sections $\sigma_s, \sigma_f, \pi_s, \pi_f$ are mapped, also as a function of neutron velocity.

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.

In the context of modelling fission in a nuclear reactor core, the dependency of the so-called cross sections $\sigma_s, \sigma_f, \pi_s, \pi_f$ on (r, v) is genuinely inhomogenous. The variable r indicates the material that neutrons pass through, which can have hugely different different interactive properties; see Figure 1. Within a given material type and at different locations in the geometry of the nuclear reactor core, the behaviour of neutrons is also highly dependent on their energy and direction of travel, which is coded by neutron velocity v .

Recent literature has explored the relationship between the NTE and a class of stochastic processes; cf. [10–12, 20, 26]. Roughly speaking, the solution to the NTE can be seen as a semigroup of the underlying physical process of neutrons, described as a branching Markov process, that is, the NBP. The NBP defined in [11, 20, 26], 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 . Analogously to (1.3) in the general MBP setting, the NBP, which we also denote $(X_t, t \geq 0)$, is given by the empirical distribution of these configurations, with $E = D \times V$ and $x_i(t) = (r_i(t), v_i(t))$.

Thus, $(X_t, t \geq 0)$ is a stochastic process valued in the space of finite counting 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_s for the random time that scattering may occur then, independently of any other physical event that may affect the neutron, $\Pr(T_s > t) = \exp\{-\int_0^t \sigma_s(r + vs, v) ds\}$, for $t \geq 0$.

When scattering occurs at space-velocity (r, v) , the new velocity $v' \in V$ is selected with probability $\pi_s(r, v, v') dv'$.

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

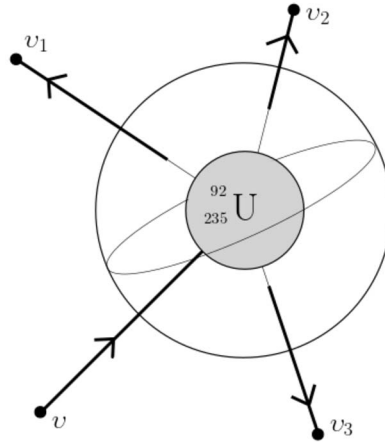


FIG. 2. A fission event between, for example, a Uranium-235 atom and a neutron with incoming velocity v produces three outgoing neutrons with velocity v_1, v_2, v_3 .

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 $\{v_i : i = 1, \dots, N\}$; see Figure 2. The outgoing velocities are described by the atomic random measure

$$(9.3) \quad \mathcal{Z}(A) := \sum_{i=1}^N \delta_{v_i}(A), \quad A \in \mathcal{B}(V).$$

If such an event occurs at location $r \in D$ from a particle with incoming velocity $v \in V$, we denote by $\mathcal{P}_{(r,v)}$ the law of \mathcal{Z} and by $\mathcal{E}_{(r,v)}$ the corresponding expectation operator. The probabilities $\mathcal{P}_{(r,v)}$ are such that, for bounded and measurable $g : V \rightarrow [0, \infty)$,

$$(9.4) \quad \int_V g(v') \pi_{\mathbb{F}}(r, v, v') dv' = \mathcal{E}_{(r,v)} \left[\int_V g(v') \mathcal{Z}(dv') \right].$$

Note, the possibility that $\Pr(N = 0) > 0$, which is tantamount to neutron capture (i.e., where a neutron slams into a nucleus but no fission results and the neutron is absorbed into the nucleus), is included in the notion of fission. Thus, the nonlocality occurs in the velocity variable. To make the connection between (9.3) and (1.4), we set $x_i = (r, v_i), i = 1, \dots, n$.

The NBP is thus parameterised by the quantities $\sigma_{\mathbb{S}}(r, v), \pi_{\mathbb{S}}(r, v, v'), \sigma_{\mathbb{F}}(r, v)$ and the measures $\mathcal{P}_{(r,v)}$ for $r \in D, v \in V$. Accordingly, we refer to it as a $(\sigma_{\mathbb{S}}, \pi_{\mathbb{S}}, \sigma_{\mathbb{F}}, \mathcal{P})$ -NBP. Although a $(\sigma_{\mathbb{S}}, \pi_{\mathbb{S}}, \sigma_{\mathbb{F}}, \mathcal{P})$ -NBP is uniquely defined, a NBP specified by $(\sigma_{\mathbb{S}}, \pi_{\mathbb{S}}, \sigma_{\mathbb{F}}, \pi_{\mathbb{F}})$ alone is not. Nonetheless, it is easy to show that, for a given $\pi_{\mathbb{F}}$, a $(\sigma_{\mathbb{S}}, \pi_{\mathbb{S}}, \sigma_{\mathbb{F}}, \mathcal{P})$ -NBP satisfying (9.4) always exists; see the discussion in Section 2 of [26].

To make the link between the NTE and the NBP clear, we will reformulate the NTE (9.1) in terms of its *Duhamel* representation. From a probabilistic perspective this is a more natural way of writing the NTE since the irregular nature of the gradient operator makes it hard to work (9.1) the pointwise sense. The following integral equation was introduced in [26] as a mild form of the NTE:

$$(9.5) \quad \psi_t[g] = \mathcal{U}_t[g] + \int_0^t \mathcal{U}_s[(S + F)\psi_{t-s}[g]] ds, \quad t \geq 0,$$

where the semigroup $(\mathcal{U}_t, t \geq 0)$ is defined by

$$\mathcal{U}_t[g](r, v) = g(r + vt, v) \mathbf{1}_{(t < \kappa_{r,v}^D)}, \quad \kappa_{r,v}^D = \inf\{t > 0 : r + vt \notin D\},$$

the operator S has action on $L^+_\infty(D \times V)$, the space of nonnegative and uniformly bounded measurable functions on $D \times V$, given by

$$Sg(r, v) = \sigma_S(r, v) \int_{D \times V} [g(r, v') - g(r, v)] \pi_S(r, v, v') dv', \quad g \in L^+_\infty(D \times V),$$

and, similarly, F has action

$$Fg(r, v) = \sigma_F(r, v) \int_{D \times V} g(r, v') \pi_F(r, v, v') dv' - \sigma_F(r, v)g(r, v), \quad g \in L^+_\infty(D \times V).$$

We encourage the reader to compare (9.5) with (2.3). Indeed, in this setting the Markov motion is just a deterministic straight line in between scattering and fission events. In particular, $P_t = U_t$ and $\mathcal{A} = S + F$. Noting $T := v \cdot \nabla_r$ as the operator corresponding to the semigroup $(U_t, t \geq 0)$, it is also possible to reformulate (9.5) as

$$\psi_t[g] = Q_t[g] + \int_0^t Q_s[F[\psi_{t-s}[g]]] ds,$$

where Q is the Markov semigroup corresponding to $T + S$. We refer the reader to [20] for further details.

It was shown in [11, 13, 14, 26] that under the assumption:²

(H1): *The cross-sections $\sigma_S, \sigma_F, \pi_S$ and π_F are uniformly bounded away from infinity*

for $g \in L^+_\infty(D \times V)$ the following linear semigroup of the NBP:

$$(9.6) \quad \psi_t[g](r, v) := \mathbb{E}_{\delta_{(r,v)}}[\langle g, X_t \rangle], \quad t \geq 0, r \in \bar{D}, v \in V$$

solves (9.5). Note that (H1) implies (G1) in this setting, since it implies that

$$m(r, v) = \mathcal{E}_{(r,v)}[\langle 1, Z \rangle] = \int_V \pi_F(r, v, v') dv'$$

is uniformly bounded from above.

As shown in [11, 26], there is a second stochastic process whose average behaviour also provides a solution to the NTE, called the neutron random walk (NRW).

Define, for $r \in D, v, v' \in V$,

$$(9.7) \quad \alpha(r, v)\pi(r, v, v') = \sigma_S(r, v)\pi_S(r, v, v') + \sigma_F(r, v)\pi_F(r, v, v').$$

Then, the $\alpha\pi$ -NRW can be defined as the process $((R_t, \Upsilon_t), t \geq 0)$ whose pathwise evolution is such that, if its current position and velocity are $R_0 = r \in D$ and $\Upsilon_0 = v \in V$, respectively, then the process will move forward in time in D with linear motion $r + vt$. Unless the NRW first exits the domain D , at rate $\alpha(r + vt, v)$, a scattering event occurs. Suppose $r' \in D$ is the NRW's position at the scattering event. Then, a new velocity v' is chosen with probability $\pi(r', v, v') dv'$. Thus, the process evolves by moving in straight lines between scatter times, until it exits the domain D . Note that, in order to specify α and π individually, it is sufficient to define the product $\alpha\pi$ since we require π to be a probability distribution.

Suppose we denote the probabilities of the $\alpha\pi$ -NRW by $\mathbf{P} := (\mathbf{P}_{(r,v)}, (r, v) \in D \times V)$. Moreover, for $r \in D$ and $v \in V$, let

$$(9.8) \quad \beta(r, v) = \sigma_F(r, v) \left(\int_V \pi_F(r, v, v') dv' - 1 \right) \geq - \sup_{r \in D, v \in V} \sigma_F(r, v) > -\infty,$$

²We note that this result was actually proved with the additional assumption that $\sigma_S\pi_S + \sigma_F\pi_F$ is uniformly bounded below on $D \times V \times V$; however, this is not actually needed.

where the lower bound is due to assumption (H1). The following lemma,³ taken from [26], identifies the law \mathbf{P} in terms of the solution to the NTE (9.5), the so-called *many-to-one* representation. Again, we encourage the reader to compare this to the general case presented in Lemma 2.1.

LEMMA 9.1 (Many-to-one). *Under assumptions (H1) the law \mathbf{P} is related to NTE (9.5) via a semigroup representation, different to (9.6), given by*

$$(9.9) \quad \psi_t[g](r, \nu) = \mathbf{E}_{(r, \nu)}[e^{\int_0^t \beta(R_s, \Upsilon_s) ds} g(R_t, \Upsilon_t) \mathbf{1}_{\{t < \tau^D\}}], \quad t \geq 0, r \in D, \nu \in V,$$

where $\tau^D = \inf\{t > 0 : R_t \notin D\}$.

As discussed in Remark 1.2, understanding the long-term growth of the NBP boils down to an understanding of the lead eigenvalue associated to the semigroup (9.6) in the spirit of Perron–Frobenius theory. In classical neutron transport theory, this corresponds to the criticality of the underlying NTE (9.1), that is, the sign of the leading eigenvalue associated to the latter.

THEOREM 9.1. *Suppose that (H1) holds, in addition to the condition:*

$$(H2): \inf_{r \in D, \nu, \nu' \in V} \alpha(r, \nu) \pi(r, \nu, \nu') > 0.$$

Then, for the semigroup $(\psi_t, t \geq 0)$ identified by (9.5), 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 $g \in L^+_\infty(D \times V)$,*

$$(9.10) \quad \langle \tilde{\varphi}, \psi_t[g] \rangle = e^{\lambda_* t} \langle \tilde{\varphi}, g \rangle \quad (\text{resp. } \psi_t[\varphi] = e^{\lambda_* t} \varphi) \quad t \geq 0.$$

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

$$(9.11) \quad \sup_{g \in L^+_{\infty,1}(D \times V)} \|e^{-\lambda_* t} \varphi^{-1} \psi_t[g] - \langle \tilde{\varphi}, g \rangle\|_\infty = O(e^{-\varepsilon t}) \quad \text{for all } t \geq 0.$$

As noted in Remark 1.2, the criticality of the NTE is determined by the sign of λ_* . When $\lambda_* > 0$ (resp., $\lambda_* < 0$), the NBP is supercritical (resp., subcritical), situations which are undesirable for nuclear reactors. The case of criticality, $\lambda_* = 0$, has the property that there is a mean steady-state for the NBP, in the sense that $\psi_t[g](r, \nu) \rightarrow \langle \tilde{\varphi}, g \rangle \varphi(r, \nu)$ as $t \rightarrow \infty$. This is the setting in which nuclear reactors typically operate and thus fuels our motivation for studying the NBP at criticality.

Despite the convergence to the steady state in the critical case, in Theorem 5.2 of [26], under the additional assumption:

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

$$\inf_{r \in B, \nu, \nu' \in V} \sigma_{\mathbb{F}}(r, \nu) \pi_{\mathbb{F}}(r, \nu, \nu') > 0,$$

³Again, this result was actually proved with the additional assumption that $\sigma_{\mathbb{S}} \pi_{\mathbb{S}} + \sigma_{\mathbb{F}} \pi_{\mathbb{F}}$ is uniformly bounded below on $D \times V \times V$; however, this is not actually needed.

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.

it was shown for the critical NBP that, for all $r \in D, v \in V$,

$$\mathbb{P}_{\delta(r,v)}(\zeta < \infty) = 1 \quad \text{where } \zeta = \inf\{t > 0 : \langle 1, X_t \rangle = 0\},$$

that is, the probability of extinction is 1.

For the rest of this section, we restrict ourselves to the case where $\lambda_* = 0$. As such, in this case Theorem 9.1 implies that (G2) holds for the NBP. Moreover, due to the above discussion, (H3) entails (G3). We will also assume that (G4) holds, which we restate in this context:

(H4): *The number of offspring produced at a fission event is bounded above by a constant N_{\max} .*

We note that (H4) is a natural physical constraint in the setting of the NBP, since fission events can only result in the release of a finite number of neutrons.

Before giving the equivalent formulations of Theorems 1.2 and 1.3 in the setting of the NBP, we will need to introduce additional notation and one more assumption. For $g \in L^+_\infty(V)$, define the spatial variance of fission velocities by

$$(9.12) \quad \mathbb{V}[g](r, v) = \mathcal{E}_{(r,v)}[\langle g, \mathcal{Z} \rangle^2 - \langle g^2, \mathcal{Z} \rangle] = \mathcal{E}_{(r,v)} \left[\sum_{\substack{1 \leq i, j \leq N \\ i \neq j}} g(v_i)g(v_j) \right].$$

We will often abuse our notation and, for $g \in L^+(D \times V)$, write $\mathbb{V}[g]$ in place of $\mathbb{V}[g(r, \cdot)]$. We will also redefine the key quantity

$$(9.13) \quad \Sigma = \langle \tilde{\varphi}, \sigma_\varepsilon \mathbb{V}[\varphi] \rangle,$$

which is finite, thanks to (H1), (H4) and the boundedness of $\tilde{\varphi}$ (cf. Theorem 9.1).

The following assumption can be seen as the NBP counterpart of (G5):

(H5): *There exists a constant $C > 0$ such that, for all $g \in L^+(D \times V)$,*

$$\langle \tilde{\varphi}, \sigma_\varepsilon \mathbb{V}[g] \rangle \geq C \langle \tilde{\varphi}, \hat{g}^2 \rangle,$$

where $\hat{g} : D \rightarrow [0, \infty) : r \mapsto \int_V g(r, v') dv'$.

As with assumption (G5), (H5) can be thought of an irreducibility type condition on the fission operator. However, in this case, when $g = u_t$, (H5) implies that if there is a fission event in the stationary distribution, the process should have a better chance of survival than producing two particles with isotropic velocities. We note that this condition is satisfied in the setting of [33] and in the isotropic case if we assume that, for all $r \in D, v \in V, \mathcal{E}_{(r,v)}[N(N - 1)] \geq c > 0$ for some constant c ,

We now restate the main results in the context of the NBP.

THEOREM 9.2. *Suppose that assumptions (H1), (H2), (H3), (H4) and (H5) hold, and $\lambda_* = 0$. Then, for all $r, \in D$ and $v \in V$,*

$$\lim_{t \rightarrow \infty} t \mathbb{P}_{\delta(r,v)}(\zeta > t) = \frac{2\varphi(r, v)}{\Sigma}.$$

THEOREM 9.3 (Yaglom limit for critical neutron transport). *Suppose that assumptions (H1), (H2), (H3), (H4) and (H5) hold with $\lambda_* = 0$. For $f \in L^+(D \times V)$ and for all $r \in D$ and $v \in V$,*

$$(9.14) \quad \lim_{t \rightarrow \infty} \mathbb{E}_{\delta(r,v)} \left[\exp\left(-\theta \frac{\langle f, X_t \rangle}{t}\right) \middle| \zeta > t \right] = \frac{1}{1 + \langle \tilde{\varphi}, f \rangle \Sigma \theta / 2}.$$

Due to the generality of our methods, the proofs of the above two theorems are almost identical to those given in the general MBP case. However, since (H5) is weaker than (G4) in the case where E is bounded, the proof of Theorem 9.2 requires some more work. More precisely, starting in the same way as in 7 and using (H5), we obtain

$$\begin{aligned}
 (9.15) \quad a'(t) &= \langle \tilde{\varphi}, \mathcal{G}[u_t] \rangle \leq -C_1 \langle \tilde{\varphi}, \sigma_{\mathbb{F}} \mathbb{V}\mathbb{V}[u_t] \rangle \\
 &\leq -C_2 \int_{D \times V} \tilde{\varphi}(r, v) \int_V \int_V u_t(r, v_1) u_t(r, v_2) \, dr \, dv \, dv_1 \, dv_2.
 \end{aligned}$$

Now, noting that Hölder’s inequality implies that

$$\left(\int_D f(r) \, dr \right)^3 \leq C_3 \int_D |f(r)|^3 \, dr,$$

we can apply this to $r \mapsto \int_V \tilde{\varphi}(r, v) u_t(r, v) \, dv$ to obtain

$$\begin{aligned}
 a(t)^3 &= \left(\int_{D \times V} \tilde{\varphi}(r, v) u_t(r, v) \, dr \, dv \right)^3 \\
 &\leq C_3 \int_D \left(\int_V \tilde{\varphi}(r, v) u_t(r, v) \, dv \right)^3 \, dr \\
 &= C_3 \int_D \, dr \int_{V \times V \times V} \tilde{\varphi}(r, v) \tilde{\varphi}(r, v_1) \tilde{\varphi}(r, v_2) u_t(r, v) u_t(r, v_1) u_t(r, v_2) \, dv \, dv_1 \, dv_2 \\
 &\leq C_4 \|u_t\|_{\infty} \int_{D \times V} \tilde{\varphi}(r, v) \int_V \int_V u_t(r, v_1) u_t(r, v_2) \, dr \, dv \, dv_1 \, dv_2,
 \end{aligned}$$

since $\tilde{\varphi}$ is uniformly bounded. Comparing this with (9.15), we have

$$a'(t) \leq -\frac{C_5}{\|u_t\|_{\infty}} a(t)^3, \quad t \geq t_0.$$

We have previously seen that $\|u_t\|_{\infty} \rightarrow 0$, which implies that, for any fixed $\epsilon > 0$, there exists $t'_0 = t'_0(\epsilon)$ so that

$$a'(t) \leq -\frac{C_5}{\epsilon^2} a(t)^3, \quad t \geq t_1 := \max(t_0, t'_0).$$

Concluding, as in Section 7, we obtain coarse upper and lower bounds but now of the order $1/\sqrt{t}$. Then, Lemma 7.4 allows us to bootstrap these bounds to obtain bounds of order $1/t$. From there, we can then conclude as in the general case.

As noted in the [Introduction](#), the work of Mori et al. [33] gives the only known results in the NBP setting. In that paper the authors took advantage of an isotropic assumption for scattering and the point process \mathcal{Z} as well as homogeneous cross-sections which allows the system to be treated as if it has local branching. More precisely, \mathcal{Z} is made up of a random number of offspring, each of which are independent and uniformly distributed on V which is otherwise taken to be the unit sphere \mathbb{S}^2 . The anisotropic behaviour of the fission we consider in this paper means that, unlike [33], we genuinely cannot treat branching as a local phenomenon and thus require more involved analysis, which may as well be delivered for general BMPs, as indeed we have done.

APPENDIX: COMBINATORIAL RESULTS

Here, we give some combinatorial results that are of importance to us. The first is the multinomial expansion formula; the second is a variant of the multinomial formula for sums of i.i.d. random variables. The third is an identity which does not appear to be easily available in the literature but is, nonetheless, relatively straightforward to prove.

LEMMA A.1 (Multinomial identity). *Suppose that a_1, \dots, a_n are real-valued and $k \in \mathbb{N}$. Then,*

$$(A.1) \quad \left(\sum_{i=1}^n a_i \right)^k = \sum_{[k_1, \dots, k_n]} \binom{k}{k_1, \dots, k_n} \prod_{i=1}^n a_i^{k_i},$$

where the sum is over the set $[k_1, \dots, k_n]$ of all combinations of nonnegative k_1, \dots, k_n such that $\sum_{i=1}^n k_i = k$ which are assigned as power indices to (a_1, \dots, a_n) .

The set $[k_1, \dots, k_n]$ is very complex, particularly when one takes into account the fact that, for example, when $n > k$, some of the k_i are necessarily zero. The combinatorics become more complicated when we apply this lemma to the setting that the a_i are i.i.d. random variables and the left-hand side of (A.1) is averaged.

LEMMA A.2. *Suppose that Y_1, \dots, Y_n are i.i.d. random variables which are equal in distribution to (Y, P) , then*

$$(A.2) \quad \mathbb{E} \left[\left(\sum_{i=1}^n Y_i \right)^k \right] = \sum_{j=1}^k \binom{n}{j} \mathbf{1}_{(j \leq n)} \sum_{[k_1, \dots, k_j]_+} \binom{k}{k_1, \dots, k_j} \prod_{i=1}^j \mathbb{E}[Y^{k_i}],$$

where the sum is over the set $[k_1, \dots, k_j]_+$ of all combinations of strictly positive k_1, \dots, k_j such that $\sum_{i=1}^j k_i = k$.

PROOF OF LEMMA A.2. We start by reexamining (A.1). To avoid having to deal with the zero values of k_i in (A.1), we would like to write

$$[k_1, \dots, k_n] = \bigcup_{j=1}^{k \wedge n} \bigcup_{\{i_1, \dots, i_j\} \in \mathcal{I}_j} \bigcup_{\{K_{i_1}, \dots, K_{i_j}\} \in [k_1, \dots, k_j]_+} \{K_1, \dots, K_n\}_{K_\ell=0, \ell \notin \{i_1, \dots, i_j\}},$$

where the first union in j identifies the number K_i from $\{K_1, \dots, K_n\} \in [k_1, \dots, k_n]$, which are strictly positive, the second union is over, \mathcal{I}_j , the set of all possible choices of j indices $\{i_1, \dots, i_j\}$ from $\{1, \dots, n\}$ and, the third union is taken over $[k_1, \dots, k_j]_+$, the set of combinations of j strictly positive integers k_1, \dots, k_j , such that $\sum_{\ell=1}^j k_\ell = k$, that are assigned to $\{K_{i_1}, \dots, K_{i_j}\}$, thereby making the element $\{K_1, \dots, K_n\}_{K_\ell=0, \ell \notin \{i_1, \dots, i_j\}}$. Note that the first union runs no further than n , as one cannot select more than n different of the Y_i -s across the k brackets in the product $(Y_1 + \dots + Y_n)^k$.

We can thus rewrite (A.1)

$$(A.3) \quad \left(\sum_{i=1}^n Y_i \right)^k = \sum_{j=1}^k \mathbf{1}_{(j \leq n)} \sum_{\{i_1, \dots, i_j\} \in \mathcal{I}_j} \sum_{[k_1, \dots, k_j]_+} \binom{k}{k_1, \dots, k_j} \prod_{\ell=1}^j Y_{i_\ell}^{k_\ell}.$$

In essence, thinking of the left-hand side as k brackets of the form $(Y_1 + \dots + Y_n)$ that are multiplied together, we first pick the indices we want to appear in multiplying one term from each bracket, then look at the number of ways these indices can be selected from each of the k brackets (this explains the appearance of the multinomial coefficient again).

The fact that Y_1, \dots, Y_n are i.i.d. means that

$$\mathbb{E} \left[\prod_{\ell=1}^j Y_{i_\ell}^{k_\ell} \right] = \prod_{\ell=1}^j \mathbb{E}[Y^{k_i}],$$

which is the same for multiple combinations of $\{i_1, \dots, i_j\}$ and $\{k_1, \dots, k_j\}$. Indeed, for a given $j \in \{1, \dots, k\}$ and combination $[k_1, \dots, k_j]_+$, there are $\binom{n}{j}$ ways of picking $j \leq n$ indices from n . In conclusion, (A.2) holds. \square

We conclude with an unusual combinatorial lemma that we could not find in the literature.

LEMMA A.3. *We have*

$$(k + 1)! = \sum_{j=1}^k 2^j \sum_{[k_1, \dots, k_j]_+} \frac{1}{j!} \binom{k}{k_1, \dots, k_j} \prod_{i=1}^j (k_i - 1)!, \quad k \in \mathbb{N}.$$

PROOF. Recall that, for $j, k \in \mathbb{N}$ such that $1 \leq j \leq k$, the (k, j) -Bell polynomial is given by

$$B_{k,j}(x_1, \dots, x_{k-j+1}) := \sum_{(\ell_1, \dots, \ell_{k-j+1})} \frac{k!}{\ell_1! \cdots \ell_{k-j+1}!} \prod_{p=1}^{k-j+1} \left(\frac{x_p}{p}\right)^{\ell_p},$$

where the sum is taken over all combinations of $(\ell_1, \dots, \ell_{k-j+1})$ such that

$$(A.4) \quad \sum_{i=1}^{k-j+1} \ell_i = j \quad \text{and} \quad \sum_{i=1}^{k-j+1} i \ell_i = k.$$

One should think of $(\ell_1, \dots, \ell_{k-j+1})$ as representing the partition of $\{1, \dots, k\}$ into $1 \leq j \leq k$ nonempty sub blocks. The quantity ℓ_i is thus the number of blocks of size i so that the first sum in (A.4) indicates that the total number of blocks is j and the total number of items in these blocks is k . Note that some of the values of ℓ_i may be zero. Note also that the index runs up to $k - j + 1$ as the largest of the j blocks can be obtained by having $j - 1$ blocks of size 1 and one block of size $k - j + 1$. As such, the (k, j) -Bell polynomial tells us the number of ways of generating j nonempty subblocks from $\{1, \dots, k\}$ by setting $x_1 = \dots = x_{k-j+1} = 1$.

If a configuration $\{k_1, \dots, k_j\}$ has block count $\{\ell_1, \dots, \ell_{k-j+1}\}$, then

$$\prod_{i=1}^j (k_i - 1)! = \prod_{p=1}^{k-j+1} ((p - 1)!)^{\ell_p}.$$

For each $j = 1, \dots, k$, counting the functionals $\prod_{i=1}^j (k_i - 1)!$ with its respective multiplicity is equivalent to counting the functionals $\prod_{p=1}^{k-j+1} ((p - 1)!)^{\ell_p}$ with their respective multiplicity.

Considering again the multinomial formula (A.2), we see that, for each $j = 1, \dots, k$ and $\{k_1, \dots, k_j\} \in [k_1, \dots, k_j]_+$, the functional $\prod_{i=1}^j (k_i - 1)!$ has a respective multiplicity given by

$$\frac{1}{j!} \binom{k}{k_1, \dots, k_j}.$$

Note, the factor $1/j!$ is present, as we are not concerned about the order of the blocks. Similarly, for each $\{\ell_1, \dots, \ell_{k-j+1}\} \in (\ell_1, \dots, \ell_{k-j+1})$, the functional $\prod_{p=1}^{k-j+1} ((p - 1)!)^{\ell_p}$ has multiplicity

$$\frac{1}{\ell_1! \cdots \ell_{k-j+1}!} \times \frac{k!}{(1!)^{\ell_1} \cdots ((k - j + 1)!)^{\ell_{k-j+1}}},$$

where the second fraction is the combinatoric for selecting ℓ_1 blocks of size 1, ℓ_2 blocks of size 2... and ℓ_{k-j+1} blocks of size $k - j + 1$, and the first fraction factors out the multiplicity of order of blocks within each size category.

It follows that:

$$\begin{aligned}
 & \sum_{j=1}^k 2^j \sum_{[k_1, \dots, k_j]_+} \frac{1}{j!} \binom{k}{k_1, \dots, k_j} \prod_{i=1}^j (k_i - 1)! \\
 &= \sum_{j=1}^k 2^j \sum_{(\ell_1, \dots, \ell_{k-j+1})} \frac{1}{\ell_1! \cdots \ell_{k-j+1}!} \frac{k!}{\prod_{p=1}^{k-j+1} (p!)^{\ell_p}} \prod_{p=1}^{k-j+1} ((p-1)!)^{\ell_p} \\
 \text{(A.5)} \quad &= \sum_{j=1}^k 2^j \sum_{(\ell_1, \dots, \ell_{k-j+1})} \frac{k!}{\ell_1! \cdots \ell_{k-j+1}!} \prod_{p=1}^{k-j+1} \left(\frac{(p-1)!}{p!} \right)^{\ell_p} \\
 &= \sum_{j=1}^k 2^j B_{k,j}(0!, \dots, (k-j)!).
 \end{aligned}$$

Standard texts on combinatorics (see, e.g., [9]) identify

$$\text{(A.6)} \quad B_{k,j}(0!, \dots, (k-j)!) = \begin{bmatrix} k \\ j \end{bmatrix},$$

which is a Stirling number of the first kind. Moreover, for the latter it is known that

$$\text{(A.7)} \quad \sum_{j=1}^k x^j \begin{bmatrix} k \\ j \end{bmatrix} = x(x+1) \cdots (x+k-1).$$

Putting the pieces together from (A.5), (A.6) and (A.7), the desired result holds. \square

Acknowledgments. We would like to thank an anonymous referee and the Associate Editor, who made a number of very helpful suggestions. We would also like to thank Ellen Powell for useful comments and our industrial partners, specifically Prof. P. Smith and Dr. G. Dobson of the ANSWERS group from Jacobs, for the use of the picture in Figure 1.

Funding. The SCH, AEK and MW were supported by EPSRC grant EP/P009220/1.

REFERENCES

[1] ATHREYA, K. B. and NEY, P. E. (2004). *Branching Processes*. Dover, Mineola, NY. MR2047480

[2] BIGGINS, J. D. (1997). How fast does a general branching random walk spread?. In *Classical and Modern Branching Processes (Minneapolis, MN, 1994)*. IMA Vol. Math. Appl. **84** 19–39. Springer, New York. MR1601689 https://doi.org/10.1007/978-1-4612-1862-3_2

[3] BILLINGSLEY, P. (1979). *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York-Chichester-Brisbane. MR0534323

[4] BREIMAN, L. (1992). *Probability. Classics in Applied Mathematics 7*. SIAM, Philadelphia, PA. MR1163370 <https://doi.org/10.1137/1.9781611971286>

[5] CHAMPAGNAT, N. and VILLEMONAIS, D. (2016). Exponential convergence to quasi-stationary distribution and Q -process. *Probab. Theory Related Fields* **164** 243–283. MR3449390 <https://doi.org/10.1007/s00440-014-0611-7>

[6] CHAMPAGNAT, N. and VILLEMONAIS, D. (2017). Uniform convergence to the Q -process. *Electron. Commun. Probab.* **22** Paper No. 33. MR3663104 <https://doi.org/10.1214/17-ECP63>

[7] CHEN, Z.-Q., REN, Y.-X. and YANG, T. (2017). Law of large numbers for branching symmetric Hunt processes with measure-valued branching rates. *J. Theoret. Probab.* **30** 898–931. MR3687244 <https://doi.org/10.1007/s10959-016-0671-y>

- [8] CHEN, Z.-Q., SONG, R. (2003). Conditional gauge theorem for non-local Feynman–Kac transforms. *Probab. Theory Related Fields* **125** 45–72. MR1952456 <https://doi.org/10.1007/s004400200219>
- [9] COMTET, L. (1974). *Advanced Combinatorics. The Art of Finite and Infinite Expansions*, enlarged ed. D. Reidel Publishing Co., Dordrecht. MR0460128
- [10] COX, A. M., HORTON, E., KYPRIANOU, A. E. and VILLEMONAIS, D. (2021). Stochastic methods for neutron transport equation III: Generational many-to-one and k_{eff} . *SIAM J. Appl. Math.* **81** 982–1001. MR4264652 <https://doi.org/10.1137/19M1295854>
- [11] COX, A. M. G., HARRIS, S. C., HORTON, E. L. and KYPRIANOU, A. E. (2019). Multi-species neutron transport equation. *J. Stat. Phys.* **176** 425–455. MR3980616 <https://doi.org/10.1007/s10955-019-02307-2>
- [12] COX, A. M. G., HARRIS, S. C., KYPRIANOU, A. E. and WANG, M. (2021). Monte carlo methods for the neutron transport equation. *SIAM J. Uncert. Quant.* Available at <https://arxiv.org/abs/2012.02864>.
- [13] DAUTRAY, R., CESSENAT, M., LEDANOIS, G., LIONS, P.-L., PARDOUX, E. and SENTIS, R. (1989). *Méthodes Probabilistes Pour les équations de la Physique. Collection du Commissariat à L'énergie Atomique*. Eyrolles, Paris.
- [14] DAUTRAY, R. and LIONS, J.-L. (1993). *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 6*. Springer, Berlin. 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. MR1295030 <https://doi.org/10.1007/978-3-642-58004-8>
- [15] DE RAPHÉLIS, L. (2017). Scaling limit of multitype Galton–Watson trees with infinitely many types. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 200–225. MR3606739 <https://doi.org/10.1214/15-AIHP713>
- [16] DYNKIN, E. B. (2002). *Diffusions, Superdiffusions and Partial Differential Equations. American Mathematical Society Colloquium Publications* **50**. Amer. Math. Soc., Providence, RI. MR1883198 <https://doi.org/10.1090/coll/050>
- [17] DYNKIN, E. B. (2004). *Superdiffusions and Positive Solutions of Nonlinear Partial Differential Equations. University Lecture Series* **34**. Amer. Math. Soc., Providence, RI. Appendix A by J.-F. Le Gall and Appendix B by I. E. Verbitsky. MR2089791 <https://doi.org/10.1090/ulect/034>
- [18] ENGLÄNDER, J., HARRIS, S. C. and KYPRIANOU, A. E. (2010). Strong law of large numbers for branching diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 279–298. MR2641779 <https://doi.org/10.1214/09-AIHP203>
- [19] GREY, D. R. (1974). Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probab.* **11** 669–677. MR0408016 <https://doi.org/10.2307/3212550>
- [20] HARRIS, S. C., HORTON, E. and KYPRIANOU, A. E. (2020). Stochastic methods for the neutron transport equation II: Almost sure growth. *Ann. Appl. Probab.* **30** 2815–2845. MR4187129 <https://doi.org/10.1214/20-AAP1574>
- [21] HARRIS, S. C. and ROBERTS, M. I. (2017). The many-to-few lemma and multiple spines. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 226–242. MR3606740 <https://doi.org/10.1214/15-AIHP714>
- [22] HARRIS, T. E. (2002). *The Theory of Branching Processes. Dover Phoenix Editions*. Dover, Mineola, NY. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)]. MR1991122
- [23] HERING, H. (1971). Critical Markov branching processes with general set of types. *Trans. Amer. Math. Soc.* **160** 185–202. MR0281272 <https://doi.org/10.2307/1995799>
- [24] HERING, H. (1973). Asymptotic behaviour of immigration-branching processes with general set of types. I. Critical branching part. *Adv. in Appl. Probab.* **5** 391–416. MR0415792 <https://doi.org/10.2307/1425827>
- [25] HERING, H. (1975). Asymptotic behaviour of immigration-branching processes with general set of types. II. Supercritical branching part. *Adv. in Appl. Probab.* **7** 468–494. MR0415793 <https://doi.org/10.2307/1426123>
- [26] HORTON, E., KYPRIANOU, A. E. and VILLEMONAIS, D. (2020). Stochastic methods for the neutron transport equation I: Linear semigroup asymptotics. *Ann. Appl. Probab.* **30** 2573–2612. MR4187122 <https://doi.org/10.1214/20-AAP1567>
- [27] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1968). Branching Markov processes. I. *J. Math. Kyoto Univ.* **8** 233–278. MR0232439 <https://doi.org/10.1215/kjm/1250524137>
- [28] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1968). Branching Markov processes. II. *J. Math. Kyoto Univ.* **8** 365–410. MR0238401 <https://doi.org/10.1215/kjm/1250524059>
- [29] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1969). Branching Markov processes. III. *J. Math. Kyoto Univ.* **9** 95–160. MR0246376 <https://doi.org/10.1215/kjm/1250524013>
- [30] KYPRIANOU, A. E. (1996). Seneta Heyde Norming in the Branching Random Walk and Associated Problems. Ph.D. thesis, Univ. Sheffield.

- [31] KYPRIANOU, A. E. (1999). A note on branching Lévy processes. *Stochastic Process. Appl.* **82** 1–14. MR1695066 [https://doi.org/10.1016/S0304-4149\(99\)00010-1](https://doi.org/10.1016/S0304-4149(99)00010-1)
- [32] KYPRIANOU, A. E. and PALAU, S. (2018). Extinction properties of multi-type continuous-state branching processes. *Stochastic Process. Appl.* **128** 3466–3489. MR3849816 <https://doi.org/10.1016/j.spa.2017.11.006>
- [33] MORI, T., WATANABE, S. and YAMADA, T. (1971/72). On neutron branching processes. *Publ. Res. Inst. Math. Sci.* **7** 153–179. MR0312584 <https://doi.org/10.2977/prims/1195193787>
- [34] MOY, S. C. (1967). Extensions of a limit theorem of Everett, Ulam and Harris on multitype branching processes to a branching process with countably many types. *Ann. Math. Stat.* **38** 992–999. MR0214155 <https://doi.org/10.1214/aoms/1177698767>
- [35] PALAU, S. and YANG, T. (2020). Law of large numbers for supercritical superprocesses with non-local branching. *Stochastic Process. Appl.* **130** 1074–1102. MR4046530 <https://doi.org/10.1016/j.spa.2019.04.007>
- [36] POWELL, E. (2019). An invariance principle for branching diffusions in bounded domains. *Probab. Theory Related Fields* **173** 999–1062. MR3936150 <https://doi.org/10.1007/s00440-018-0847-8>
- [37] REN, Y.-X., SONG, R. and SUN, Z. (2018). A 2-spine decomposition of the critical Galton–Watson tree and a probabilistic proof of Yaglom’s theorem. *Electron. Commun. Probab.* **23** Paper No. 42. MR3841403 <https://doi.org/10.1214/18-ECP143>
- [38] REN, Y.-X., SONG, R. and SUN, Z. (2020). Limit theorems for a class of critical superprocesses with stable branching. *Stochastic Process. Appl.* **130** 4358–4391. MR4102269 <https://doi.org/10.1016/j.spa.2020.01.001>
- [39] REN, Y.-X., SONG, R. and SUN, Z. (2020). Spine decompositions and limit theorems for a class of critical superprocesses. *Acta Appl. Math.* **165** 91–131. MR4058118 <https://doi.org/10.1007/s10440-019-00243-7>
- [40] SEVAST’YANOV, B. A. (1958). Branching stochastic processes for particles diffusing in a bounded domain with absorbing boundaries. *Theory Probab. Appl.* **3** 111–126.
- [41] UCHIYAMA, K. (1982). Spatial growth of a branching process of particles living in \mathbf{R}^d . *Ann. Probab.* **10** 896–918. MR0672291