



HAL
open science

Asymptotic moments of spatial branching processes

Isaac Gonzalez, Emma Horton, Andreas E Kyprianou

► **To cite this version:**

Isaac Gonzalez, Emma Horton, Andreas E Kyprianou. Asymptotic moments of spatial branching processes. *Probability Theory and Related Fields*, 2022, 184, pp.805-858. 10.1007/s00440-022-01131-2. hal-03321803v2

HAL Id: hal-03321803

<https://hal.science/hal-03321803v2>

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic moments of spatial branching processes

Isaac Gonzalez^{*,†}, Emma Horton[‡] and Andreas E. Kyprianou^{*}

December 20, 2021

Abstract

Suppose that $X = (X_t, t \geq 0)$ is either a superprocess or a branching Markov process on a general space E , with non-local branching mechanism and probabilities \mathbb{P}_{δ_x} , when issued from a unit mass at $x \in E$. For a general setting in which the first moment semigroup of X displays a Perron-Frobenius type behaviour, we show that, for $k \geq 2$ and any positive bounded measurable function f on E ,

$$\lim_{t \rightarrow \infty} g(t) \mathbb{E}_{\delta_x} [\langle f, X_t \rangle^k] = C_k(x, f),$$

where the constant $C_k(x, f)$ can be identified in terms of the principal right eigenfunction and left eigen-measure and $g(t)$ is an appropriate deterministic normalisation, which can be identified explicitly as either polynomial in t or exponential in t , depending on whether X is a critical, supercritical or subcritical process. The method we employ is extremely robust and we are able to extract similarly precise results that additionally give us the moment growth with time of $\int_0^t \langle g, X_s \rangle ds$, for bounded measurable g on E .

Key words: Moments, branching processes, superprocesses, non-local branching, asymptotic behaviour

MSC 2020: 60J68, 60J80.

1 Introduction and main results

A fundamental question concerning general spatial branching processes, both superprocesses and branching Markov processes, pertains to their moments. Whilst the setting of first and second moments has received quite some attention, limited information seems to be known about higher moments, in particular, their asymptotic behaviour with time. Relevant references that touch upon this topic include [11, 7, 15, 18, 12]. In this paper, we provide a single general result that pertains to both superprocesses and spatial branching Markov processes and which provides a very precise and somewhat remarkable result for moment growth.

^{*}Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK. E-mails: igg22@bath.ac.uk, a.kyprianou@bath.ac.uk

[†]Research supported by CONACYT scholarship nr 472301

[‡]INRIA, Bordeaux Research Centre, 33405 Talence, France. E-mail: emma.horton@inria.fr

We show that, under the assumption that the first moment semigroup of the process exhibits a natural Perron Frobenius type behaviour, the k -th moment functional of either a superprocess or branching Markov process, when appropriately normalised, limits to a precise constant. The setting in which we work is remarkably general, even allowing for the setting of non-local branching; that is, where mass is created at a different point in space to the position of the parent. Moreover, the methodology we use appears to be extremely robust and we show that the asymptotic k -th moments of the running occupation measure are equally accessible using essentially the same approach. Our results will thus expand on what is known for branching diffusions and superdiffusions e.g. in [5], [17], as well as giving precise growth rates for the moments of occupations.

To this end, let us spend some time providing the general setting in which we wish to work. Let E be a Lusin space. Throughout, will write $B(E)$ for the Banach space of bounded measurable functions on E with norm $\|\cdot\|$, $B^+(E)$ for non-negative bounded measurable functions on E and $B_1^+(E)$ for the subset of functions in $B^+(E)$ which are uniformly bounded by unity. We are interested in spatial branching processes that are defined in terms of a Markov process and a branching operator. The former can be characterised by a semigroup on E , denoted by $\mathbf{P} = (\mathbf{P}_t, t \geq 0)$. We do not need \mathbf{P} to have the Feller property, and it is not necessary that \mathbf{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 \mathbf{P} as conservative on the extended space $E \cup \{\dagger\}$, which can also be treated as a Lusin space. Equally, we can extend the branching operator to $E \cup \{\dagger\}$ by defining it to be zero on $\{\dagger\}$, i.e. no branching activity on the cemetery state.

1.1 Branching Markov processes

Consider now a spatial branching process in which, given their point of creation, particles evolve independently according to a \mathbf{P} -Markov process. In an event which we refer to as ‘branching’, particles positioned at x die at rate $\beta(x)$, where $\beta \in B^+(E)$, and instantaneously, new particles are created in E according to a point process. The configurations of these offspring are described by the random counting measure

$$\mathcal{Z}(A) = \sum_{i=1}^N \delta_{x_i}(A),$$

for Borel A in E . The law of the aforementioned point process depends on x , the point of death of the parent, and we denote it by \mathcal{P}_x , $x \in E$, with associated expectation operator given by \mathcal{E}_x , $x \in E$. This information is captured in the so-called branching mechanism

$$(1) \quad \mathbf{G}[f](x) := \beta(x) \mathcal{E}_x \left[\prod_{i=1}^N f(x_i) - f(x) \right], \quad x \in E,$$

where we recall $f \in B_1^+(E) := \{f \in B^+(E) : \sup_{x \in E} f(x) \leq 1\}$. Without loss of generality we can assume that $\mathcal{P}_x(N = 1) = 0$ for all $x \in E$ by viewing a branching event with one offspring as an extra jump in the motion. On the other hand, we do allow for the possibility that $\mathcal{P}_x(N = 0) > 0$ for some or all $x \in E$.

Henceforth we refer to this spatial branching process as a (\mathbf{P}, \mathbf{G}) -branching Markov process. It is well known that if the configuration of particles at time t is denoted by $\{x_1(t), \dots, x_{N_t}(t)\}$, then, on the event that the process has not become extinct or exploded, the branching Markov process can be described as the co-ordinate process $X = (X_t, t \geq 0)$ in the space of atomic measures on E with non-negative integer total mass, denoted by $N(E)$, where

$$X_t(\cdot) = \sum_{i=1}^{N_t} \delta_{x_i(t)}(\cdot), \quad t \geq 0.$$

In particular, X is Markovian in $N(E)$. Its probabilities will be denoted $\mathbb{P} := (\mathbb{P}_\mu, \mu \in N(E))$. With this notation in hand, it is worth noting that the independence that is manifest in the definition of branching events and movement implies that if we define,

$$(2) \quad \mathbf{v}_t[f](x) = \mathbb{E}_{\delta_x} \left[\prod_{i=1}^{N_t} f(x_i(t)) \right], \quad f \in B_1^+(E), t \geq 0,$$

then for $\mu \in N(E)$ given by $\mu = \sum_{i=1}^n \delta_{y_i}$, we have

$$(3) \quad \mathbb{E}_\mu \left[\prod_{i=1}^{N_t} f(x_i(t)) \right] = \prod_{i=1}^n \mathbf{v}_t[f](y_i), \quad t \geq 0.$$

Moreover, for $f \in B^+(E)$ and $x \in E$,

$$(4) \quad \mathbf{v}_t[f](x) = \hat{\mathbf{P}}_t[f](x) + \int_0^t \mathbf{P}_s[\mathbf{G}[\mathbf{v}_{t-s}[f]]](x) ds, \quad t \geq 0,$$

where $\hat{\mathbf{P}}_t$ is a slight adjustment of \mathbf{P}_t which returns a value of 1 on the event of killing.

Branching Markov processes enjoy a very long history in the literature, dating back as far as [23, 22, 24], with a broad base of literature that is arguably too voluminous to give a fair summary of here. Most literature focuses on the setting of local branching. This corresponds to the setting that all offspring are positioned at their parent's point of death (i.e. $x_i = x$ in the definition of \mathbf{G}). In that case, the branching mechanism reduces to

$$\mathbf{G}[s](x) = \beta(x) \left[\sum_{k=1}^{\infty} p_k(x) s^k - s \right], \quad x \in E,$$

where $s \in [0, 1]$ and $(p_k(x), k \geq 0)$ is the offspring distribution when a parent branches at site $x \in E$. The branching mechanism \mathbf{G} may otherwise be seen in general as a mixture of local and non-local branching.

1.2 Superprocesses

Superprocesses can be thought of as the high-density limit of a sequence of branching Markov processes, resulting in a new family of measure-valued Markov processes; see e.g. [19, 3, 25,

6, 4]. Just as branching Markov processes are Markovian in $N(E)$, the former are Markovian in the space of finite Borel measures on E topologised by the weak convergence topology, denoted by $M(E)$. There is a broad literature base for superprocesses, e.g. [19, 3, 25, 11, 9], with so-called local branching mechanisms, and later broadened to the more general setting of non-local branching mechanisms in [4, 19]. Let us now introduce these concepts with an autonomous definition of what we mean by a superprocess.

A Markov process $X := (X_t : t \geq 0)$ with state space $M(E)$ and probabilities $\mathbb{P} := (\mathbb{P}_\mu, \mu \in M(E))$ is called a (\mathbb{P}, ψ, ϕ) -superprocess if it has transition semigroup $(\hat{\mathbf{E}}_t, t \geq 0)$ on $M(E)$ satisfying

$$(5) \quad \mathbb{E}_\mu [e^{-\langle f, X_t \rangle}] = \int_{M(E)} e^{-\langle f, \nu \rangle} \hat{\mathbf{E}}_t(\mu, d\nu) = e^{-\langle \mathbf{V}_t[f], \mu \rangle}, \quad \mu \in M(E), f \in B^+(E).$$

Here, we work with the inner product on $B^+(E) \times M(E)$ defined by $\langle f, \mu \rangle = \int_E f(x) \mu(dx)$ and $(\mathbf{V}_t, t \geq 0)$ is a semigroup evolution that is characterised via the unique bounded positive solution to the evolution equation

$$(6) \quad \mathbf{V}_t[f](x) = \mathbf{P}_t[f](x) - \int_0^t \mathbf{P}_s [\psi(\cdot, \mathbf{V}_{t-s}[f](\cdot)) + \phi(\cdot, \mathbf{V}_{t-s}[f])] (x) ds.$$

Here ψ denotes the local branching mechanism

$$(7) \quad \psi(x, \lambda) = -b(x)\lambda + c(x)\lambda^2 + \int_{(0, \infty)} (e^{-\lambda y} - 1 + \lambda y) \nu(x, dy), \quad \lambda \geq 0,$$

where $b \in B(E)$, $c \in B^+(E)$ and $(x \wedge x^2)\nu(x, dy)$ is a bounded kernel from E to $(0, \infty)$, and ϕ is the non-local branching mechanism

$$(8) \quad \phi(x, f) = \beta(x) (f(x) - \zeta(x, f)),$$

where $\beta \in B^+(E)$ and ζ has representation

$$(9) \quad \zeta(x, f) = \gamma(x, f) + \int_{M(E)^\circ} (1 - e^{-\langle f, \nu \rangle}) \Gamma(x, d\nu),$$

such that $\gamma(x, f)$ is a bounded function on $E \times B^+(E)$ and $\nu(1)\Gamma(x, d\nu)$ is a bounded kernel from E to $M(E)^\circ := M(E) \setminus \{0\}$ with

$$(10) \quad \gamma(x, f) + \int_{M(E)^\circ} \langle 1, \nu \rangle \Gamma(x, d\nu) \leq 1.$$

Lemma 3.1 in [4] tells us that the functional $\zeta(x, f)$ has the following equivalent representation

$$(11) \quad \zeta(x, f) = \int_{M_0(E)} \left[\gamma(x, \pi) \langle f, \pi \rangle + \int_0^\infty (1 - e^{-u \langle f, \pi \rangle}) n(x, \pi, du) \right] G(x, d\pi),$$

where $M_0(E)$ denotes the set of probability measures on E , $d \geq 0$ is a bounded function on $E \times M_0(E)$, $un(x, \pi, du)$ is a bounded kernel from $E \times M_0(E)$ to $(0, \infty)$ and $G(x, d\pi)$ is a probability kernel from E to $M_0(E)$ with

$$(12) \quad \gamma(x, \pi) + \int_0^\infty un(x, \pi, du) \leq 1.$$

The reader will note that we have deliberately used some of the same notation for both branching Markov processes and superprocesses. In the sequel there should be no confusion and the motivation for this choice of repeated notation is that our main result is indifferent to which of the two processes we are talking about.

1.3 Main results: k -th moments

As alluded to above, in what follows, (X, \mathbb{P}) is taken as either a branching Markov process or a superprocess as defined in the previous section. Our main results concern understanding the growth of the k -th moment functional in time

$$\mathbf{T}_t^{(k)}[f](x) := \mathbb{E}_{\delta_x} \left[\langle f, X_t \rangle^k \right], \quad x \in E, f \in B^+(E), k \geq 1, t \geq 0.$$

For convenience, we will write \mathbf{T} in preference of $\mathbf{T}^{(1)}$ throughout.

Before stating our main theorem, we first introduce some assumptions that will be crucial in analysing the moments defined above. First, we have a Perron-Frobenius-type assumption.

(H1): There exists an eigenvalue $\lambda \in \mathbb{R}$ and a corresponding right eigenfunction $\varphi \in B^+(E)$ and finite left eigenmeasure $\tilde{\varphi}$ such that, for $f \in B^+(E)$,

$$\langle \mathbf{T}_t[\varphi], \mu \rangle = e^{\lambda t} \langle \varphi, \mu \rangle \text{ and } \langle \mathbf{T}_t[f], \tilde{\varphi} \rangle = e^{\lambda t} \langle f, \tilde{\varphi} \rangle,$$

for all $\mu \in N(E)$ (resp. $M(E)$) if (X, \mathbb{P}) is a branching Markov process (resp. a superprocess). Further let us define

$$\Delta_t = \sup_{x \in E, f \in B_1^+(E)} |\varphi(x)^{-1} e^{-\lambda t} \mathbf{T}_t[f](x) - \langle \tilde{\varphi}, f \rangle|, \quad t \geq 0.$$

We suppose that

$$(13) \quad \sup_{t \geq 0} \Delta_t < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t = 0.$$

The second assumption is a moment condition on the offspring distribution.

(H2): Suppose $k \geq 1$. If (X, \mathbb{P}) is a branching Markov process,

$$(14) \quad \sup_{x \in E} \mathcal{E}_x(\langle 1, \mathcal{Z} \rangle^k) < \infty$$

and if (X, \mathbb{P}) is a superprocess,

$$(15) \quad \sup_{x \in E} \left(\int_0^\infty |y|^k \nu(x, dy) + \int_{M(E)^\circ} \langle 1, \nu \rangle^k \Gamma(x, d\nu) \right) < \infty.$$

Let us spend a little time considering these two assumptions in more detail. 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 as in (H1); see e.g. [21, 10, 1, 20, 16, 14]. Due to this, we may characterise the process as supercritical if $\lambda > 0$, critical if $\lambda = 0$ and subcritical if $\lambda < 0$.

One way to understand (13), is through the martingale that comes hand-in-hand with the eigenpair (λ, φ) , i.e.

$$(16) \quad M_t^\varphi := e^{-\lambda t} \langle \varphi, X_t \rangle, \quad t \geq 0.$$

Normalising this martingale and using it as a change of measure results in the ubiquitous spine decomposition; cf. [16, 15, 21]. Roughly speaking, under the change of measure, the process is equal in law to a copy of the original process with a superimposed process of immigration, which occurs both in space and time along the path of a single particle trajectory in E , the spine. If the process is issued from e.g. $\mu \in M(E)$, then the semigroup of the latter, is given by $e^{-\lambda t} \langle \mathbb{T}_t[f\varphi], \mu \rangle / \langle \varphi, \mu \rangle$, $t \geq 0$. We see then that an assumption of the type (13) implies that the spine has a stationary limit with stationary measure $\varphi \tilde{\varphi}$.

The assumptions (14) and (15) of (H2) are natural to ensure that k -moments are well defined for all $t \geq 0$. If not explicitly stated in the literature, their need to ensure that the functional moments $\mathbb{T}_t^{(k)}[f](x)$ are finite for all $t \geq 0$, $f \in B^+(E)$ and $x \in E$ is certainly folklore. The two conditions (14) and (15) are clearly natural analogues of one another. Indeed, whereas for superprocesses, it is usual to separate out the non-diffusive local branching behaviour from non-local behaviour, i.e via the measures $\nu(x, dy)$ and $\Gamma(x, d\nu)$, the analogous behaviour is captured in the single point process \mathcal{Z} for branching Markov processes.

In terms of the eigenvalue in (H1), the following suite of results give us the precise growth rates for k -th moments in each of the critical ($\lambda = 0$), supercritical ($\lambda > 0$) and subcritical ($\lambda < 0$) settings. In all three results, (X, \mathbb{P}) is either a (\mathbb{P}, \mathbb{G}) -branching particle system or a (\mathbb{P}, ψ, ϕ) -superprocess on E .

Theorem 1 (Critical, $\lambda = 0$). *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda = 0$. Define*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| t^{-(\ell-1)} \varphi(x)^{-1} \mathbb{T}_t^{(\ell)}[f](x) - 2^{-(\ell-1)} \ell! \langle f, \tilde{\varphi} \rangle^\ell \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{\ell-1} \right|,$$

where

$$\mathbb{V}[\varphi](x) = \beta(x) \mathcal{E}_x \left(\langle \varphi, \mathcal{Z} \rangle^2 - \langle \varphi^2, \mathcal{Z} \rangle \right),$$

if (X, \mathbb{P}) is a branching Markov process or

$$\mathbb{V}[\varphi](x) = \psi''(x, 0+) \varphi(x)^2 + \beta(x) \int_{M(E)^\circ} \langle \varphi, \nu \rangle^2 \Gamma(x, d\nu)$$

if (X, \mathbb{P}) is a superprocess. Then, for all $\ell \leq k$

$$(17) \quad \sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t^{(\ell)} = 0.$$

The novel contribution of Theorem 1 is both the fact that no such result currently exists in the literature as well as the general and precise polynomial growth of the k -th moment. The only other comparable result is that of [13], which inspired this paper and deals with the special case of a general *critical* branching particle processes and that the test function f is specifically taken to be the eigenfunction φ . There are two facts that stand out in this result. The first is the polynomial scaling, which is quite a delicate conclusion given that there is no exponential growth to rely on. The second is that, for $k \geq 3$, the scaled moment limit is expressed not in terms of the k -th moments in (14) and (15), but rather the second order moments.

In some sense, however, both the polynomial growth and the nature of the limiting constant are not entirely surprising given the folklore for the critical setting. More precisely, in at least some settings (see e.g. [13]), one would expect to see a Yaglom-type result at criticality. The latter would classically see convergence of $t^{-1} \langle f, X_t \rangle$ in law to an exponentially distributed random variable as $t \rightarrow \infty$, whose parameter is entirely determined by the second moments of X .

The next results present a significantly different picture for the supercritical and subcritical cases. For those settings, the exponential behaviour of the first moment semigroup becomes a dominant feature of the higher expected moments.

Theorem 2 (Supercritical, $\lambda > 0$). *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda > 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-\ell \lambda t} \mathbf{T}_t^{(\ell)}[f](x) - \ell! \langle f, \tilde{\varphi} \rangle^\ell L_\ell \right|,$$

where $L_1 = 1$ and we define iteratively for $k \geq 2$

$$L_k = \frac{1}{\lambda(k-1)} \left\langle \beta \mathcal{E} \left[\sum_{[k_1, \dots, k_N]_k^2} \prod_{j=1}^N \varphi(x_j) L_{k_j} \right], \tilde{\varphi} \right\rangle,$$

where $[k_1, \dots, k_N]_k^2$ is the set of all non-negative N -tuples (k_1, \dots, k_N) such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive¹ if (X, \mathbb{P}) is a branching Markov process, or

$$L_k(x) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{1}{m_1! \dots m_{k-1}!} (m_1 + \dots + m_{k-1} - 1)! \varphi(x)^{m_1 + \dots + m_{k-1} - 1} \prod_{j=1}^{k-1} (-L_j(x))^{m_j} \\ + \frac{1}{\lambda(k-1)} \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle$$

¹We interpret $\sum_\emptyset = 0$ and $\prod_\emptyset = 1$.

and iteratively with $\mathbb{V}_2[\varphi](x) = \mathbb{V}[\varphi](x)$ (defined in the previous theorem) and for $k \geq 3$

$$\begin{aligned} & \mathbb{V}_k[\varphi](x) \\ &= \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{1}{m_1! \dots m_{k-1}!} \prod_{j=2}^{k-1} \left(\frac{\langle \mathbb{V}_j[\varphi], \tilde{\varphi} \rangle}{\lambda(j-1)} \right)^{m_j} \\ & \quad \times \left[\psi^{(m_1 + \dots + m_{k-1})}(x, 0+) (-\varphi(x))^{m_1 + \dots + m_{k-1}} + \beta(x) \int_{M(E)^\circ} \langle \varphi, \nu \rangle^{m_1 + \dots + m_{k-1}} \Gamma(x, d\nu) \right] \end{aligned}$$

if (X, \mathbb{P}) is a superprocess. Here the sums run over the set $\{m_1, \dots, m_{k-1}\}_k$ of positive integers such that $m_1 + 2m_2 + \dots + (k-1)m_{k-1} = k$. Then, for all $\ell \leq k$

$$(18) \quad \sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t^{(\ell)} = 0.$$

As with the critical setting, we could not find any existing result of this kind in the literature. While Jensen's inequality easily shows that this is the minimal rate of growth, it turns out that it is the exact rate of growth. If we again appeal to folklore then this is again not necessarily surprising. In a number of settings, we would expect X to obey a strong law of large numbers (cf. [14, 10, 1, 20]) in the sense that

$$\lim_{t \rightarrow \infty} e^{-\lambda t} \langle f, X_t \rangle = \langle \tilde{\varphi}, f \rangle M_\infty^\varphi,$$

where $(M_t^\varphi, t \geq 0)$ was defined in (16) and the limit holds either almost surely or in the sense of L^p moments, for $p > 1$. It is also worth remarking on the fact that there is dependency on x for the limit in the superprocess setting but not in the branching Markov process setting. Thinking in terms of the skeletal decomposition (see e.g. [8, 2]), in principle the superprocess issued from a unit mass at x can be seen as the aggregation of a Poisson point process of 'superprocess excursions'. In the supercritical setting, a finite Poisson number of these will contribute to the overall growth of the process which are sampled at a rate proportional to $p(x)\delta_x$, where $p(x)$ is rate of survival of an excursion issued from $x \in E$. This may go part way to explaining the dependency of L_k on x in that setting.

Finally we turn to the growth of moments in the subcritical setting, which offers the heuristically appealing result that the k -th moment decays slower than the k -th moment of the linear semigroup.

Theorem 3 (Subcritical, $\lambda < 0$). *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda < 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-\lambda t} \mathbf{T}_t^{(\ell)}[f](x) - \ell! \langle f, \tilde{\varphi} \rangle^\ell L_\ell \right|,$$

where we define iteratively $L_1 = 1$ and for $k \geq 2$,

$$L_k = \frac{\langle f^k, \tilde{\varphi} \rangle}{\langle f, \tilde{\varphi} \rangle^k k!} + \left\langle \beta \mathcal{E} \cdot \left[\sum_{n=2}^k \frac{1}{|\lambda|(n-1)} \sum_{[k_1, \dots, k_N]_k^n} \prod_{\substack{j=1 \\ j:k_j > 0}}^N \varphi(x_j) L_{k_j} \right], \tilde{\varphi} \right\rangle,$$

where $[k_1, \dots, k_N]_k^n$ is the set of all non-negative N -tuples (k_1, \dots, k_N) such that $\sum_{i=1}^N k_i = k$ and exactly $2 \leq n \leq k$ of the k_i are strictly positive if (X, \mathbb{P}) is a branching Markov process, or $L_k = \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle$, where $\mathbb{V}_1[\varphi](x) = \varphi(x)$ and for $k \geq 2$,

$$\begin{aligned} \mathbb{V}_k[\varphi](x) = & \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{1}{m_1! \dots m_{k-1}!} \frac{1}{\lambda(1 - m_1 - \dots - m_{k-1})} \prod_{j=1}^{k-1} \langle \mathbb{V}_j[\varphi], \tilde{\varphi} \rangle^{m_j} \\ & \times \left[\psi^{(m_1 + \dots + m_{k-1})}(x, 0+) (-\varphi(x))^{m_1 + \dots + m_{k-1}} + \beta(x) \int_{M(E)^\circ} \langle \varphi, \nu \rangle^{m_1 + \dots + m_{k-1}} \Gamma(x, d\nu) \right] \end{aligned}$$

if (X, \mathbb{P}) is a superprocess. Here the sums run over the set $\{m_1, \dots, m_{k-1}\}_k$ of positive integers such that $m_1 + 2m_2 + \dots + (k-1)m_{k-1} = k$. Then, for all $\ell \leq k$

$$(19) \quad \sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t^{(\ell)} = 0.$$

As alluded to above, it is heuristically appealing that the the k -th moment does not grow at the rate $\exp(-k\lambda t)$. On the other hand the actual growth rate $\exp(-\lambda t)$ is slightly less obvious but nonetheless the obvious candidate. The decay in mass to zero in the branching system would suggest that the k -th moment similarly does so, but no slower than the first moment.

The robustness of our methods in the following sections means that the principal ideas used to prove the above theorems are essentially the same for both branching particle systems and superprocesses, regardless of the criticality. The main idea is to study the non-linear semigroup, $\mathbb{E}_{\delta_x}[e^{-\theta \langle f, X_t \rangle}]$, $\theta, t \geq 0, f \in B^+(E)$, associated with (X, \mathbb{P}) . The relation

$$(20) \quad \mathbf{T}^{(k)}[f](x) = (-1)^k \frac{\partial}{\partial \theta} \mathbb{E}_{\delta_x}[e^{-\theta \langle f, X_t \rangle}] \Big|_{\theta=0}$$

then means that we can use knowledge of the non-linear semigroup to study the moments of (X, \mathbb{P}) . Indeed, the first step is to write an evolution equation for $\mathbb{E}_{\delta_x}[e^{-\theta \langle f, X_t \rangle}]$ in terms of the linear semigroup \mathbf{T} . We will show that differentiating this equation and using (20) means that we can write the k -th moment of the process in terms of the lower order moments. An inductive argument will then yield the final results.

Despite this generic approach, the proof for each of the two processes requires slightly different technicalities due to the fact that, on the one hand, superprocesses have Lévy-Khintchine branching mechanisms but no particles, whereas on the other hand, branching particle systems do have individual particles but less regular branching mechanisms. Due to these discrepancies, we require a different toolbox to deal with the evolution equation of the k -th moment. To compute the derivatives in (20) for the branching particle systems, we use the Leibniz rule, however, in the case of superprocesses, we use Fàa di Bruno's rule. This yields different equations for the k -th moment evolutions, resulting in slightly different combinatorial arguments when completing the proofs of the theorems. Moreover, the different normalisation required in each of the three theorems, as well as the different limits requires some care. For this reason, we will prove Theorem 1 in detail, since this case

requires the most delicate analysis, and in the remaining non-critical cases, we will give an outline of the proofs, leaving the details to the reader.

It also transpires that our method is remarkably robust. Indeed, again taking an agnostic position on whether X is a branching Markov process or a superprocess, as we will show, careful consideration of the proofs of Theorems 1, 2 and 3 demonstrate that we can also conclude results for the quantities

$$\mathbf{M}_t^{(k)}[g](x) := \mathbb{E}_{\delta_x} \left[\left(\int_0^t \langle g, X_s \rangle ds \right)^k \right], \quad x \in E, g \in B^+(E), k \geq 1, t \geq 0.$$

We can think of $\int_0^t \langle g, X_s \rangle ds$ as characterising the running occupation measure $\int_0^t X_s(\cdot) ds$ of the process X and hence we refer to $\mathbf{M}_t^{(k)}[g](x)$ as the k -th moment of the running occupation. The following results also emerge from our calculations, mirroring Theorems 1, 2 and 3 respectively.

Theorem 4 (Critical, $\lambda = 0$). *Suppose that (H1) holds along with (H2) for $k \geq 2$ and $\lambda = 0$. Define*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| t^{-(2\ell-1)} \varphi(x)^{-1} \mathbf{M}_t^{(\ell)}[g](x) - 2^{-(\ell-1)} \ell! \langle g, \tilde{\varphi} \rangle^\ell \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{\ell-1} L_\ell \right|,$$

where $L_1 = 1$ and L_k is defined through the recursion $L_k = (\sum_{i=1}^{k-1} L_i L_{k-i}) / (2k-1)$ if (X, \mathbb{P}) is a branching Markov process or $L_k = (\sum_{\{k_1, k_2\}^+} L_{k_1} L_{k_2}) / (2k-1)$ where $\{k_1, k_2\}^+$ is the set of positive integers k_1, k_2 such that $k_1 + k_2 = k$ if (X, \mathbb{P}) is a superprocess. Then, for all $\ell \leq k$

$$(21) \quad \sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t^{(\ell)} = 0.$$

Theorem 5 (Supercritical, $\lambda > 0$). *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda > 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-\ell \lambda t} \mathbf{M}_t^{(\ell)}[g](x) - \ell! \langle g, \tilde{\varphi} \rangle^\ell L_\ell \right|,$$

where L_k was defined in Theorem 2, albeit that $L_1 = 1/\lambda$.

Then, for all $\ell \leq k$

$$(22) \quad \sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t^{(\ell)} = 0.$$

Theorem 6 (Subcritical, $\lambda < 0$). *Suppose that (H1) holds along with (H2) for some $k \geq 2$ and $\lambda < 0$. Redefine*

$$\Delta_t^{(\ell)} = \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} \mathbf{M}_t^{(\ell)}[g](x) - \ell! \langle g, \tilde{\varphi} \rangle^\ell L_\ell \right|,$$

where $L_1 = 1/|\lambda|$ and for $k \geq 2$, the constants L_k are defined recursively via

$$L_k = \frac{1}{|\lambda|} \left\langle \beta \mathcal{E} \left[\sum_{[k_1, \dots, k_N]_k} \prod_{\substack{j=1 \\ j:k_j > 0}}^N \varphi(x_j) L_{k_j} \right], \tilde{\varphi} \right\rangle - \frac{\langle g\varphi, \tilde{\varphi} \rangle}{|\lambda| \langle g, \tilde{\varphi} \rangle} L_{k-1},$$

if X is a branching Markov process and

$$L_k(x) = (-1)^k \langle \mathbb{U}_k[\varphi], \tilde{\varphi} \rangle - \mathcal{R}_k(x),$$

where

$$\mathcal{R}_k(x) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{(-1)^k}{m_1! \dots m_{k-1}!} (-\varphi(x))^{m_1 + \dots + m_{k-1} - 1} (m_1 + \dots + m_{k-1} - 1)! \prod_{j=1}^{k-1} L_j(x)^{m_j},$$

and where we define recursively, $\mathbb{U}_1[\varphi](x) = -\varphi(x)/|\lambda|$ and for $k \geq 2$,

$$\begin{aligned} & \mathbb{U}_k[\varphi](x) \\ &= \frac{1}{|\lambda|} \left(\mathbb{U}_{k-1}[\varphi](x) + \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{1}{m_1! \dots m_{k-1}!} \prod_{j=1}^{k-1} \langle \mathbb{U}_j[\varphi], \tilde{\varphi} \rangle^{m_j} \right. \\ & \quad \left. \times \left[\psi^{(m_1 + \dots + m_{k-1})}(x, 0+) (-\varphi(x))^{m_1 + \dots + m_{k-1}} + \beta(x) \int_{M(E)^\circ} \langle \varphi, \nu \rangle^{m_1 + \dots + m_{k-1}} \Gamma(x, d\nu) \right] \right) \end{aligned}$$

if X is a superprocess. Then, for all $\ell \leq k$

$$(23) \quad \sup_{t \geq 0} \Delta_t^{(\ell)} < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t^{(\ell)} = 0.$$

The results in Theorems 4, 5 and 6 are slightly less predictable. Let us discuss this point a little further in the particle setting for convenience. For the supercritical case, the extra ‘‘linear’’ term arising from the time integral does not affect the exponential growth of the process, and hence the leading order behaviour is still dominated by $e^{k\lambda t}$. In the critical case, we know from Theorem 1 that the first moment does not require normalisation, and hence integrating up to time t will induce a linear growth in time. As we will shortly see, one can determine the k -th moment from a combination of the lower order moments and as such, this linear growth will propagate through the recursion, which along with the time integral, yields the t^{2k-1} scaling. Finally, in the subcritical case, we know that the total occupation $\int_0^\zeta \langle g, X_s \rangle ds$, where $\zeta = \inf\{t > 0 : \langle 1, X_t \rangle = 0\}$, is finite, behaving like an average spatial distribution of mass, i.e. $\langle g, \tilde{\varphi} \rangle$, multiplied by ζ , meaning that no normalisation is required to control the ‘‘growth’’ of the running occupation moments in this case.

The rest of this paper can be summarised as follows. In the next section, we focus on the case where (X, \mathbb{P}) is a branching particle process. We first prove Theorem 1, followed by an outline of the proofs of Theorems 2 and 3. We then turn to the case where (X, \mathbb{P}) is a

superprocess in section 3. Following a similar format, we first prove the result for the critical case, followed by a sketch of the proofs for the non-critical cases. As indicated above, both approaches follow a similar pattern and thus both sections are laid out to reflect this. First of all, we study the linear and non-linear evolution equations associated with (X, \mathbb{P}) . The relation (20) then yields an evolution equation for the k -th moment in terms of the lower order moments. Using this and an inductive argument, along with several crucial results that we house in the appendix, yields the result of the three theorems.

2 Proofs for branching Markov processes

The proof is a mixture of analytical and combinatorial computations which are based around the behaviour of the linear and non-linear semigroups of X .

2.1 Linear and non-linear semigroup equations

For $f \in B^+(E)$, it is well known that the mean semigroup evolution satisfies

$$(24) \quad \mathbb{T}_t[f](x) = \mathbb{P}_t[f] + \int_0^t \mathbb{P}_s [\mathbb{F}\mathbb{T}_{t-s}[f]](x) ds \quad t \geq 0, x \in E,$$

where

$$\mathbb{F}[f](x) = \beta(x) \mathcal{E}_x \left[\sum_{i=1}^N f(x_i) - f(x) \right] =: \beta(x) (\mathbf{m}[f](x) - f(x)), \quad x \in E.$$

See for example the calculations in [16]. Associated with every linear semigroup of a branching process is a so-called many-to-one formula. Many-to-one formulae are not necessarily unique and the one we will develop here is slightly different from the usual construction because of non-locality.

Suppose that $\xi = (\xi_t, t \geq 0)$, with probabilities $\mathbf{P} = (\mathbf{P}_x, x \in E)$, is the Markov process corresponding to the semigroup \mathbb{P} . Let us introduce a new Markov process $\hat{\xi} = (\hat{\xi}_t, t \geq 0)$ which evolves as the process ξ but at rate $\beta(x)\mathbf{m}[1](x)$ the process is sent to a new position in E , such that for all Borel $A \subset E$, the new position is in A with probability $\mathbf{m}[1_A](x)/\mathbf{m}[1](x)$. We will refer to the latter as *extra jumps*. Note the law of the extra jumps is well defined thanks to the assumption (14), which ensures that $\sup_{x \in E} \mathbf{m}[1](x) = \sup_{x \in E} \mathcal{E}_x(\langle 1, \mathcal{Z} \rangle) < \infty$. Accordingly we denote the probabilities of $\hat{\xi}$ by $(\hat{\mathbf{P}}_x, x \in E)$. We can now state our many-to-one formula.

Lemma 1. *Write $B(x) = \beta(x)(\mathbf{m}[1](x) - 1)$, $x \in E$. For $f \in B^+(E)$ and $t \geq 0$, we have*

$$(25) \quad \mathbb{T}_t[f](x) = \hat{\mathbf{E}}_x \left[\exp \left(\int_0^t B(\hat{\xi}_s) ds \right) f(\hat{\xi}_t) \right].$$

The proof is classical and follows standard reasoning for semigroup integral equations e.g. as in [16, 14]: First conditioning the right-hand side of (25) on the time of the first extra

jump, then using the principle of transferring between multiplicative and additive potentials in the resulting integral equation (cf. Lemma 1.2, Chapter 4 in [6]) shows that (24) holds. Grönwall's Lemma, the fact that $\beta \in B^+(E)$ and (14) for $k = 1$ ensure that the relevant integral equations have unique solutions.

We now define a variant of the non-linear evolution equation (2) associated with X via

$$(26) \quad \mathbf{u}_t[f, g](x) = \mathbb{E}_{\delta_x} \left[1 - e^{-\langle f, X_t \rangle - \int_0^t \langle g, X_s \rangle ds} \right], \quad t \geq 0, x \in E, f, g \in B^+(E).$$

For $f \in B_1^+(E)$, define

$$\mathbf{A}[f](x) = \beta(x) \mathcal{E}_x \left[\prod_{i=1}^N (1 - f(x_i)) - 1 + \sum_{i=1}^N f(x_i) \right], \quad x \in E.$$

Our first preparatory result relates the two semigroups $(\mathbf{u}_t, t \geq 0)$ and $(\mathbf{T}_t, t \geq 0)$.

Lemma 2. *For all $f, g \in B^+(E)$, $x \in E$ and $t \geq 0$, the non-linear semigroup $\mathbf{u}_t[f, g](x)$ satisfies*

$$(27) \quad \mathbf{u}_t[f, g](x) = \mathbf{T}_t[1 - e^{-f}](x) - \int_0^t \mathbf{T}_s [\mathbf{A}[\mathbf{u}_{t-s}[f, g]] - g(1 - \mathbf{u}_{t-s}[f, g])](x) ds.$$

Proof. Again, the proof uses standard techniques for integral evolution equations so we only sketch the proof. Instead of considering $\mathbf{u}_t[f, g]$, we will first work instead with

$$(28) \quad \mathbf{v}_t[f, g] = \mathbb{E}_{\delta_x} \left[e^{-\langle f, X_t \rangle - \int_0^t \langle g, X_s \rangle ds} \right], \quad t \geq 0, x \in E, f, g \in B^+(E),$$

which will turn out to be more convenient for technical reasons.

By splitting the expectation in (28) on the first branching event and appealing to the Markov property, we get, for $f, g \in B^+(E)$, $t \geq 0$ and $x \in E$,

$$\begin{aligned} \mathbf{v}_t[f, g](x) &= \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_s) ds} e^{-f(\xi_t) - \int_0^t g(\xi_s) ds} \right] \\ &\quad + \mathbf{E}_x \left[\int_0^t \beta(\xi_s) e^{-\int_0^s \beta(\xi_u) + g(\xi_u) du} \mathbf{H}[\mathbf{v}_{t-s}[f, g]](\xi_s) ds \right], \end{aligned}$$

where

$$\mathbf{H}[g](x) = \mathcal{E}_x \left[\prod_{i=1}^N g(x_i) \right], \quad g \in B^+(E), x \in E.$$

Using similar reasoning to Lemma 1.2, Chapter 4 in [6] we can move the multiplicative potential with rate $\beta + g$ to an additive potential in the above evolution equation to obtain

$$(29) \quad \mathbf{v}_t[f, g](x) = \hat{\mathbf{P}}_t[e^{-f}](x) + \int_0^t \mathbf{P}_s [\mathbf{G}[\mathbf{v}_{t-s}[f, g]] - g\mathbf{v}_{t-s}[f, g]](x) ds.$$

Now define

$$\mathbf{D}[f](x) = \beta(x) \mathcal{E}_x \left[\prod_{i=1}^N f(x_i) - \sum_{i=1}^N f(x_i) \right] = \beta(x) (\mathbf{H}[f](x) - \mathbf{m}[f](x)), \quad f \in B_1^+(E), x \in E$$

and $(\tilde{v}_t, t \geq 0)$ via

$$\begin{aligned}
\tilde{v}_t[f, g](x) &= \mathbf{T}_t[e^{-f}](x) + \int_0^t \mathbf{T}_s \left[\mathbf{D}[\tilde{v}_{t-s}[f, g]] - g\tilde{v}_{t-s}[f, g] \right](x) ds \\
&= \hat{\mathbf{E}}_x \left[e^{\int_0^t \mathbf{B}(\hat{\xi}_s) ds} e^{-f(\hat{\xi}_t)} \right] \\
(30) \quad &+ \hat{\mathbf{E}}_x \left[\int_0^t e^{\int_0^s \mathbf{B}(\hat{\xi}_u) du} \left(\mathbf{D}[\tilde{v}_{t-s}[f, g]](\hat{\xi}_s) - g(\hat{\xi}_s)\tilde{v}_{t-s}[f, g](\hat{\xi}_s) \right) ds \right],
\end{aligned}$$

for $x \in E, t \geq 0$ and $f, g \in B^+(E)$. Note that for the moment we don't claim a solution to (30) exists.

For convenience, we will define

$$\mathbf{K}_t[f, g](x) = \hat{\mathbf{E}}_x \left[\int_0^t e^{\int_0^s \mathbf{B}(\hat{\xi}_u) du} \left(\mathbf{D}[\tilde{v}_{t-s}[f, g]](\hat{\xi}_s) - g(\hat{\xi}_s)\tilde{v}_{t-s}[f, g](\hat{\xi}_s) \right) ds \right],$$

so that $\tilde{v}_t[f, g](x) = \mathbf{T}_t[e^{-f}](x) + \mathbf{K}_t[f, g](x)$. By conditioning the right-hand side of (30) on the first jump of $\hat{\xi}$ (bearing in mind the dynamics of $\hat{\xi}$ given just before Lemma 1) with the help of the Markov property (recalling that $\mathbf{B}(x) - \beta \mathbf{m}[1] = \beta$), we get

$$\begin{aligned}
&\tilde{v}_t[f, g](x) \\
&= \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_s) ds} e^{-f(\xi_t)} \right] + \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell) \mathbf{m}[1](\xi_\ell) e^{-\int_0^\ell \beta(\xi_s) ds} \frac{\mathbf{m}[\mathbf{T}_{t-\ell}[e^{-f}]](\xi_\ell)}{\mathbf{m}[1](\xi_\ell)} d\ell \right] \\
&+ \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_u) \mathbf{m}[1](\xi_u) du} \int_0^t e^{\int_0^s \mathbf{B}(\xi_u) du} \left(\mathbf{D}[\tilde{v}_{t-s}[f, g]](\xi_s) - g(\xi_s)\tilde{v}_{t-s}[f, g](\xi_s) \right) ds \right] \\
&+ \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell) \mathbf{m}[1](\xi_\ell) e^{-\int_0^\ell \beta(\xi_u) \mathbf{m}[1](\xi_u) du} \right. \\
&\quad \left. \left(\int_0^\ell e^{\int_0^s \mathbf{B}(\xi_u) du} \left(\mathbf{D}[\tilde{v}_{t-s}[f, g]](\xi_s) - g(\xi_s)\tilde{v}_{t-s}[f, g](\xi_s) \right) ds \right. \right. \\
&\quad \left. \left. + e^{\int_0^\ell \mathbf{B}(\xi_u) du} \frac{\mathbf{m}[\mathbf{K}_{t-\ell}[g]](\xi_\ell)}{\mathbf{m}[1](\xi_\ell)} \right) d\ell \right].
\end{aligned}$$

Gathering terms and exchanging the order of integration in the double integral, this simplifies

to

$$\begin{aligned}
& \tilde{\mathbf{v}}_t[f, g](x) \\
&= \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_s) ds} e^{-f(\xi_t)} \right] + \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell) e^{-\int_0^\ell \beta(\xi_s) ds} \mathbf{m}[\tilde{\mathbf{v}}_{t-\ell}[f, g](x)](\xi_\ell) d\ell \right] \\
& \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_u) \mathbf{m}[1](\xi_u) du} \int_0^t e^{\int_0^s \mathbf{B}(\xi_u) du} \left(\mathbf{D}[\tilde{\mathbf{v}}_{t-s}[f, g]](\xi_s) - g(\xi_s) [\tilde{\mathbf{v}}_{t-s}[f, g]] \right) ds \right] \\
& + \mathbf{E}_x \left[\int_0^t \int_0^t \mathbf{1}_{(s \leq \ell)} \beta(\xi_\ell) \mathbf{m}[1](\xi_\ell) e^{-\int_0^\ell \beta(\xi_u) \mathbf{m}[1](\xi_u) du} e^{\int_0^s \mathbf{B}(\xi_u) du} \right. \\
& \quad \left. \left(\mathbf{D}[\tilde{\mathbf{v}}_{t-s}[f, g]](\xi_s) - g(\xi_s) \tilde{\mathbf{v}}_{t-s}[f, g](\xi_s) \right) d\ell ds \right] \\
&= \mathbf{E}_x \left[e^{-\int_0^t \beta(\xi_s) ds} e^{-f(\xi_t)} \right] + \mathbf{E}_x \left[\int_0^t \beta(\xi_\ell) e^{-\int_0^\ell \beta(\xi_s) ds} \mathbf{m}[\tilde{\mathbf{v}}_{t-\ell}[g](x)](\xi_\ell) d\ell \right] \\
& + \mathbf{E}_x \left[\int_0^t e^{-\int_0^s \beta(\xi_u) du} \left(\mathbf{D}[\tilde{\mathbf{v}}_{t-s}[f, g]](\xi_s) - g(\xi_s) \tilde{\mathbf{v}}_{t-s}[f, g](\xi_s) \right) ds \right].
\end{aligned}$$

Finally, appealing to the change of multiplicative potential to additive potential in the spirit of e.g. Lemma 1.2, Chapter 4 of [6], we get

$$\tilde{\mathbf{v}}_t[f, g](x) = \hat{\mathbf{P}}_t[e^{-f}](x) + \int_0^t \mathbf{P}_t \left[\mathbf{G}[\tilde{\mathbf{v}}_{t-s}[f, g]] - g \tilde{\mathbf{v}}_{t-s}[f, g] \right](x) ds$$

and hence $(\tilde{\mathbf{v}}_t, t \geq 0)$ is a solution to (29). A standard argument using using $\beta \in B^+(E)$, the assumption (14) for $k = 1$ and Grönwall's Lemma tells us that all of the integral equations thus far have unique solutions. In conclusion, $(\mathbf{v}_t[g], t \geq 0)$ and $(\tilde{\mathbf{v}}_t[g], t \geq 0)$ agree.

To complete the lemma, note that

$$1 - \mathbf{T}_t[e^{-f}](x) = \mathbf{T}_t[1 - e^{-f}](x) + 1 - \mathbf{T}_t[1](x)$$

moreover,

$$1 - \mathbf{T}_t[1](x) = \hat{\mathbf{E}}_x \left[\int_0^t \mathbf{B}(\hat{\xi}_s) e^{\int_0^s \mathbf{B}(\hat{\xi}_u) du} ds \right] = \int_0^t \mathbf{T}_s[\mathbf{B}] ds.$$

Hence, working from (30) and the definitions of \mathbf{D} and \mathbf{A} , which are related via

$$\mathbf{D}[1 - f] = \beta(x) \mathcal{E}_x \left[\prod_i (1 - f(x_i)) - \sum_{i=1}^N (1 - f(x_i)) \right] = \mathbf{A}[f] + \mathbf{B}(x), \quad x \in E, f \in B_1^+(E),$$

we get

$$\begin{aligned}
\mathbf{u}_t[f, g](x) &= 1 - \mathbf{v}_t[f, g](x) \\
&= 1 - \mathbf{T}_t[e^{-f}](x) - \int_0^t \mathbf{T}_s \left[\mathbf{D}[1 - \mathbf{u}_{t-s}[f, g]] - g(1 - \mathbf{u}_{t-s}[f, g]) \right](x) ds \\
&= \mathbf{T}_t[1 - e^{-f}] - \int_0^t \mathbf{T}_s \left[\mathbf{A}[\mathbf{u}_{t-s}[f, g]] - g(1 - \mathbf{u}_{t-s}[f, g]) \right](x) ds,
\end{aligned}$$

as required. \square

2.2 Evolution equations for the k -th moment of branching Markov processes

Next we turn our attention to the evolution equation generated by the k -th moment functional $\mathbf{T}_t^{(k)}$, $t \geq 0$. To this end, we start by observing that

$$(31) \quad \mathbf{T}_t^{(k)}[f](x) = (-1)^{k+1} \frac{\partial^k}{\partial \theta^k} \mathbf{u}_t[e^{-\theta f}](x) \Big|_{\theta=0}.$$

The following result gives us an iterative approach to writing the k -th moment functional in terms of lower order moment functionals.

Proposition 1. *Fix $k \geq 2$. Under the assumptions of Theorem 1, with the additional assumption that*

$$(32) \quad \sup_{x \in E, s \leq t} \mathbf{T}_s^{(\ell)}[f](x) < \infty, \quad \ell \leq k-1, f \in B^+(E), t \geq 0,$$

it holds that

$$(33) \quad \mathbf{T}_t^{(k)}[f](x) = \mathbf{T}_t[f^k](x) + \int_0^t \mathbf{T}_s \left[\beta \eta_{t-s}^{(k-1)}[f] \right](x) ds, \quad t \geq 0,$$

where

$$\eta_{t-s}^{(k-1)}[f](x) = \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathbf{T}_{t-s}^{(k_j)}[f](x_j) \right],$$

and $[k_1, \dots, k_N]_k^2$ is the set of all non-negative N -tuples (k_1, \dots, k_N) such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive.

Proof. Recall from (27) that

$$(34) \quad \mathbf{u}_t[\theta f, 0](x) = \mathbf{T}_t[1 - e^{-\theta f}](x) - \int_0^t \mathbf{T}_s [\mathbf{A}[\mathbf{u}_{t-s}[\theta f, 0]]](x) ds, \quad t \geq 0.$$

It is clear that differentiating the first term k times and setting $\theta = 0$ on the right-hand side of (34) yields

$$(35) \quad \frac{\partial^k}{\partial \theta^k} \mathbf{T}_t[1 - e^{-\theta f}](x) \Big|_{\theta=0} = (-1)^{k+1} \mathbf{T}[f^k](x).$$

Thus it remains to differentiate the second term on the right-hand side of (34) k times. To this end, without concern for passing derivatives through expectations, using the Leibniz

rule in Lemma A.2 of the Appendix, we have

$$\begin{aligned}
& -\frac{\partial^k}{\partial \theta^k} \mathbf{A}[\mathbf{u}_t[\theta f, 0]](x) \Big|_{\theta=0} \\
&= \frac{\partial^k}{\partial \theta^k} \beta(x) \mathcal{E}_x \left[1 - \prod_{i=1}^N \mathbb{E}_{\delta_{x_i}} [e^{-\theta \langle f, X_i \rangle}] - \sum_{i=1}^N \mathbb{E}_{\delta_{x_i}} [1 - e^{-\theta \langle f, X_i \rangle}] \right] \\
&= -\beta(x) \mathcal{E}_x \left[\sum_{k_1 + \dots + k_N = k} \binom{k}{k_1, \dots, k_m} \prod_{j=1}^N (-1)^{k_j} \mathbf{T}_t^{(k_j)}[f](x_j) + (-1)^{k+1} \sum_{i=1}^N \mathbf{T}_t^{(k)}[f](x_i) \right] \\
(36) \quad &= \beta(x) \mathcal{E}_x \left[(-1)^{k+1} \sum_{k_1 + \dots + k_N = k} \binom{k}{k_1, \dots, k_m} \prod_{j=1}^N \mathbf{T}_t^{(k_j)}[f](x_j) + (-1)^k \sum_{i=1}^N \mathbf{T}_t^{(k)}[f](x_i) \right].
\end{aligned}$$

where the sum is taken over all non-negative integers k_1, \dots, k_N such that $\sum_{i=1}^N k_i = k$.

Next let us look in more detail at the sum/product term on the righthand (36). Consider the terms where only one of the k_i in the sum is positive, in which case $k_i = k$ and

$$\binom{k}{k_1, \dots, k_m} = 1.$$

There are N ways this can happen in the sum of the sum-product term and hence

$$\begin{aligned}
& \sum_{k_1 + \dots + k_N = k} \binom{k}{k_1, \dots, k_m} \prod_{j=1}^N \mathbf{T}_t^{(k_j)}[f](x_j) \\
&= \sum_{i=1}^N \mathbf{T}_t^{(k)}[f](x_i) + \sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathbf{T}_t^{(k_j)}[f](x_j),
\end{aligned}$$

where $[k_1, \dots, k_N]_k^2$ is the set of all non-negative N -tuples (k_1, \dots, k_N) such that $\sum_{i=1}^N k_i = k$ and at least two of the k_i are strictly positive. Substituting this back into (36) yields

$$-\frac{\partial^k}{\partial \theta^k} \mathbf{A}[\mathbf{u}_t[e^{-\theta f}]] \Big|_{\theta=0} = (-1)^{k+1} \beta(x) \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathbf{T}_t^{(k_j)}[f](x_j) \right].$$

Now let us return to the justification that we can pass the derivatives through the expectation in the above calculation, we first note that derivatives are limits and so an ‘epsilon-delta’ argument will ultimately require dominated convergence. This is where the assumption (14) and (32) come in. On the right-hand side of (36), each of the $\mathbf{T}_t^{(k_j)}[f](x_j)$ in the sum term are uniformly bounded by the assumption (32) and the collection $[k_1, \dots, k_N]_k^2$ means that $0 \leq k_j \leq k - 1$ for each $j = 1, \dots, N$. Moreover, there can be at most k items in the sum/product. Noting that

$$(37) \quad \sum_{k_1 + \dots + k_N = k} \binom{k}{k_1, \dots, k_m} = N^k,$$

the assumption (14) allows us to use a domination argument with the k -th order moment.

Combining this with (35) and (34), using an easy dominated convergence argument to pull the k derivatives through the integral in t , then dividing by $(-1)^{k+1}$, we get (33), as required. \square

2.3 Completing the proof of Theorem 1: critical case

We will prove Theorem 1 by induction, starting with the case $k = 1$. In this case, (21) reads

$$\sup_{t \geq 0} \Delta_t < \infty \text{ and } \lim_{t \rightarrow \infty} \Delta_t = 0,$$

which holds due to (13).

We now assume that the theorem holds true in the branching Markov process setting for some $k \geq 1$ and proceed to show that (21) holds for all $\ell \leq k + 1$.

To this end, first note that the induction hypothesis implies that (32) holds. Hence Proposition 1 tells us that

$$\begin{aligned} & \varphi(x)^{-1} t^{-k} \mathbf{T}_t^{(k+1)}[f](x) \\ &= \varphi(x)^{-1} t^{-k} \mathbf{T}_t[f^{(k+1)}](x) \\ & \quad + \varphi(x)^{-1} t^{-k} \int_0^t \mathbf{T}_s \left[\mathcal{E} \cdot \left[\sum_{[k_1, \dots, k_N]_{k+1}^2} \binom{k+1}{k_1, \dots, k_N} \prod_{j=1}^N \mathbf{T}_{t-s}^{(k_j)}[f](x_j) \right] \right] (x) ds \\ &= \varphi(x)^{-1} t^{-k} \mathbf{T}_t[f^{(k+1)}](x) \\ (38) \quad & \quad + \varphi(x)^{-1} t^{-(k-1)} \int_0^1 \mathbf{T}_{ut} \left[\mathcal{E} \cdot \left[\sum_{[k_1, \dots, k_N]_{k+1}^2} \binom{k+1}{k_1, \dots, k_N} \prod_{j=1}^N \mathbf{T}_{t(1-u)}^{(k_j)}[f](x_j) \right] \right] (x) du, \end{aligned}$$

where we have used the change of variables $s = ut$ in the final equality.

We now make some observations that will simplify the expression on the right-hand side of (38) as $t \rightarrow \infty$. First note that due to (13), the first term on the righthand side of (38) will vanish as $t \rightarrow \infty$. Next, note that, if more than two of the k_i in the sum are strictly positive, then the renormalising by t^{k-1} will cause the associated summand to go to zero as well. For example, suppose without loss of generality that k_1 and k_2 are both strictly positive, we can write $t^{k-1} = t^{(k+1)-2} = t^{k_1-1} t^{k_2-1} t^{k_3} \dots t^{k_N}$. Now the induction hypothesis tells us that the correct normalisation of each of the terms in the product is t^{k_j-1} , which means that the item $\mathbf{T}_{t(1-u)}^{(k_j)}$ for a third $k_j > 0$ will be ‘over normalised to zero’ in the limit.

To make this heuristic rigorous, we can employ Theorem A.1 from the Appendix. To this end, let us set

$$(39) \quad F[f](x, u, t) := \frac{1}{\varphi(x) t^{k-1}} \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_{k+1}^3} \binom{k+1}{k_1, \dots, k_N} \prod_{j=1}^N \mathbf{T}_{t(1-u)}^{(k_j)}[f](x_j) \right]$$

where $[k_1, \dots, k_N]_{k+1}^3$ is the subset of $[k_1, \dots, k_N]_{k+1}^2$, for which at least three of the k_i are strictly positive (which can be an empty set). We will show that conditions (A.1) and (A.2) are satisfied via

$$(40) \quad \sup_{x \in E, f \in B_1^+(E), u \in [0,1]} \varphi(x)F[f](x, u, t) < \infty \text{ and } \lim_{t \rightarrow \infty} \sup_{u \in [0,1], f \in B_1^+(E), x \in E} \varphi(x)F[f](x, u, t) = 0.$$

First note that there are no more than $k + 1$ of the k_i that are strictly greater than 1 in the product in (39). This follows from the fact that it is not possible to partition the set $\{1, \dots, k + 1\}$ into more than $k + 1$ non-empty blocks. Next note that

$$\frac{1}{t^{k-1}} \prod_{\substack{j=1 \\ j:k_j > 0}}^N \mathbf{T}_{t(1-u)}^{(k_j)}[f](x_j) = \frac{(t(1-u))^{k+1-\#\{j:k_j > 0\}}}{t^{k-1}} \prod_{\substack{j=1 \\ j:k_j > 0}}^N \varphi(x_j) \cdot \frac{1}{\varphi(x_j)} \frac{\mathbf{T}_{t(1-u)}^{(k_j)}[f](x_j)}{(t(1-u))^{k_j-1}}.$$

The product term on the right-hand side is uniformly bounded in x_j and $t(1-u)$ on compact intervals due to boundedness of φ and the fact that (21) is assumed to hold for all $\ell \leq k$ by induction. Moreover, if $\#\{j : k_j > 0\} \leq 1$, the set $[k_1, \dots, k_N]_{k+1}^3$ is empty, otherwise, the term $(t(1-u))^{k+1-\#\{j:k_j > 0\}}/t^{k-1}$ is finite for all $t \geq 1$, say. From (37) and (14), we also observe that

$$\sup_{x \in E} \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_{k+1}^3} \binom{k+1}{k_1, \dots, k_N} \right] \leq \sup_{x \in E} \mathcal{E}_x [\langle 1, \mathcal{Z} \rangle^{k+1}] < \infty.$$

Taking these facts into account, it is now straightforward to see that the earlier given heuristic can be made rigorous and (40) holds. In particular, we can use dominated convergence to pass the limit in t through the expectation in (39) to achieve the second statement in (40).

As F belongs to the class of functions \mathcal{C} , defined just before Theorem A.1 in the Appendix, the aforesaid theorem tells us that

$$(41) \quad \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \int_0^1 \mathbf{T}_{ut}[\varphi F[f](\cdot, u, t)](x) du \right| = 0.$$

Returning to (38), since the sum there requires that at least two of the k_i are positive, this means that the only surviving terms in the limit are those that are combinations of two strictly positive terms k_i and k_j such that $i \neq j$ and $k_i + k_j = k + 1$. This can be thought of as choosing $i, j \in \{1, \dots, N\}$ with $i \neq j$, choosing $k_i \in \{1, \dots, k\}$ and then setting $k_j = k + 1 - k_i$. One should take care however to avoid double counting each pair (k_i, k_j) . Thus, we have

$$(42) \quad \frac{1}{t^k \varphi(x)} \mathbf{T}_t^{(k+1)}[f](x) = \frac{1}{\varphi(x)} \int_0^1 \mathbf{T}_{ut} \left[\frac{\beta(\cdot)}{2t^{k-1}} \mathcal{E} \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^k \binom{k+1}{k_i, k+1-k_i} \right. \right. \\ \left. \left. \times \mathbf{T}_{t(1-u)}^{(k_i)}[f](x_i) \mathbf{T}_{t(1-u)}^{(k+1-k_i)}[f](x_j) \right] \right] (x) du,$$

where the factor of $1/2$ appears to compensate for the aforementioned double counting.

In order to show that the right-hand side above delivers the required finiteness and limit (21), we again turn to Theorem A.1. For $x \in E$, $t \geq 0$ and $0 \leq u \leq 1$, in anticipation of using this theorem, we now re-define

$$F[f](x, u, t) := \frac{\beta(x)}{2\varphi(x)t^{k-1}} \mathcal{E}_x \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^k \binom{k+1}{k_i, k+1-k_i} \mathbb{T}_{t(1-u)}^{(k_i)}[f](x_i) \mathbb{T}_{t(1-u)}^{(k+1-k_i)}[f](x_j) \right].$$

After some rearrangement, we have

$$\begin{aligned} & F[f](x, u, t) \\ &= \frac{\beta(x)(1-u)^{k-1}}{2\varphi(x)} \mathcal{E}_x \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^k \binom{k+1}{k_i, k+1-k_i} \right. \\ & \quad \left. \times \varphi(x_i)\varphi(x_j) \frac{\mathbb{T}_{t(1-u)}^{(k_i)}[f](x_i)}{\varphi(x_i)(t(1-u))^{k_i-1}} \frac{\mathbb{T}_{t(1-u)}^{(k+1-k_i)}[f](x_j)}{\varphi(x_j)(t(1-u))^{k-k_i}} \right]. \end{aligned} \quad (43)$$

Using similar arguments to those given previously in the proof of (41) may, again, combine the induction hypothesis, simple combinatorics and dominated convergence to pass the limit as $t \rightarrow \infty$ through the expectation and show that

$$\begin{aligned} & F[f](x, u) := \lim_{t \rightarrow \infty} F[f](x, u, t) \\ &= (k+1)! \langle \tilde{\varphi}, \mathbb{V}[\varphi] \rangle / 2)^{k-1} \langle \tilde{\varphi}, f \rangle^{k+1} k \frac{(1-u)^{k-1}}{2\varphi(x)} \mathbb{V}[\varphi](x), \end{aligned} \quad (44)$$

for which one uses that

$$\begin{aligned} & (k+1)! \langle \tilde{\varphi}, \beta \mathbb{V}[\varphi] \rangle / 2)^{k-1} \langle \tilde{\varphi}, f \rangle^{k+1} k \mathbb{V}[\varphi](x) \\ &= \mathcal{E}_x \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^k \binom{k+1}{k_i, k+1-k_i} \varphi(x_i)\varphi(x_j) \right. \\ & \quad \left. \times \frac{k_i! \langle f, \tilde{\varphi} \rangle^{k_i} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k_i-1}}{2^{(k_i-1)}} \frac{(k+1-k_i)! \langle f, \tilde{\varphi} \rangle^{k+1-k_i} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-k_i}}{2^{(k-k_i)}} \right]. \end{aligned}$$

Note that, thanks to the assumption (H2), the expression for $F(s, x)$ clearly satisfies (A.1).

Subtracting the right-hand side of (44) from the right-hand side of (43), again appealing to the induction hypotheses, specifically the second statement in (21), it is not difficult to show that, for each $\varepsilon \in (0, 1)$,

$$\lim_{t \rightarrow \infty} \sup_{x \in E, u \in [0, \varepsilon], f \in B_1^+(E)} |\varphi(x)F[f](x, u, t) - \varphi(x)F[f](x, u)| = 0.$$

On the other hand, the first statement in the induction hypothesis (21) also implies that there exists a constant $C_k > 0$ (which depends on k but not ε) such that

$$\lim_{t \rightarrow \infty} \sup_{x \in E, u \in [\varepsilon, 1], f \in B_1^+(E)} |\varphi(x)F(x, u, t) - \varphi(x)F(x, u)| \leq C_k(1 - \varepsilon)^{k-1}.$$

Since we may take ε arbitrarily close to 1, we conclude that (A.2) holds.

In conclusion, since the conditions of Theorem A.1 are now met, we get the two statements of (21) as a consequence. \square

2.4 Proof of Theorem 4

Next we turn our attention to the evolution equation generated by the k -th moment functional $\mathbf{T}_t^{(k)}$, $t \geq 0$. To this end, we start by defining observing that

$$(45) \quad \mathbf{M}_t^{(k)}[g](x) = (-1)^{k+1} \frac{\partial^k}{\partial \theta^k} \mathbf{u}_t[0, \theta g](x) \Big|_{\theta=0}.$$

Taking account of (27), we see that

$$(46) \quad \mathbf{u}_t[0, \theta g](x) = - \int_0^t \mathbf{T}_s [\mathbf{A}[\mathbf{u}_{t-s}[0, \theta g]] - \theta g(1 - \mathbf{u}_{t-s}[0, \theta g])] (x) ds.$$

Given the proximity of (46) to (34), it is easy to see that we can apply the same reasoning that we used for $\mathbf{T}_t^{(k)}[f](x)$ to $\mathbf{M}_t^{(k)}[g](x)$ and conclude that, for $k \geq 2$,

$$(47) \quad \mathbf{M}_t^{(k)}[g](x) = \int_0^t \mathbf{T}_s [\beta \hat{\eta}_{t-s}^{(k-1)}[g]](x) - k \mathbf{T}_s [g \mathbf{M}_{t-s}^{(k-1)}[g]](x) ds,$$

where $\hat{\eta}^k$ plays the role of η^k albeit replacing the moment operators $\mathbf{T}^{(j)}$ by the moment operators $\mathbf{M}^{(j)}$.

We now proceed to prove Theorem 4, also by induction. First we consider the setting $k = 1$. In that case,

$$\frac{1}{t} \mathbf{M}^{(1)}[g](x) = \frac{1}{t} \mathbb{E}_{\delta_x} \left[\int_0^t \langle g, X_s \rangle ds \right] = \frac{1}{t} \int_0^t \mathbf{T}_s [g](x) ds = \int_0^1 \mathbf{T}_{ut} [g](x) du.$$

Referring now to Theorem A.1 in the Appendix, we can take $F(x, s, t) = f(x)/\varphi(x)$, since $f \in B^+(E)$, the conditions of the theorem are trivially met and hence

$$\lim_{t \rightarrow \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{t} \mathbf{M}^{(1)}[g](x) - \langle g, \tilde{\varphi} \rangle \right| = 0.$$

Note that this limit sets the scene for the polynomial growth in $t^{n(k)}$ of the higher moments for some function $n(k)$. If we are to argue by induction, whatever the choice of $n(k)$, it must satisfy $n(1) = 1$.

Next suppose that Theorem 4 holds for all integer moments up to and including $k - 1$. We have from (47) that

$$(48) \quad \frac{1}{t^{2k-1}} \mathbf{M}_t^{(k)}[g](x) = \frac{1}{t^{2k-1}} \int_0^t \mathbf{T}_s \left[\beta \hat{\eta}_{t-s}^{(k-1)}[g] \right](x) ds - \frac{1}{t^{2k-1}} \int_0^t k \mathbf{T}_s [g \mathbf{M}_{t-s}^{(k-1)}[g]](x) ds.$$

Let us first deal with the right most integral in (48). It can be written as

$$\frac{1}{t^{2k-2}} \int_0^1 k \mathbf{T}_{ut} [\varphi F(\cdot, u, t)](x) du := \int_0^1 (1-u)^{2k-2} k \mathbf{T}_{ut} \left[g \frac{1}{(t(1-u))^{2k-2}} \mathbf{M}_{t(1-u)}^{(k-1)}[g] \right](x) du.$$

Arguing as in the spirit of the proof of Theorem 1, our induction hypothesis ensures that

$$\lim_{t \rightarrow \infty} F[g](x, u, t) = \lim_{t \rightarrow \infty} g(1-u)^{2k-2} k \frac{1}{(t(1-u))^{2k-2}} \frac{\mathbf{M}_{t(1-u)}^{(k-1)}[g](x)}{\varphi(x)} = 0 =: F(x, u)$$

satisfies (A.1) and (A.2). Theorem A.1 thus tells us that, uniformly in $x \in E$ and $g \in B_1^+(E)$,

$$(49) \quad \lim_{t \rightarrow \infty} \frac{1}{t^{2k-1}} \int_0^t k \mathbf{T}_s [g \mathbf{M}_{t-s}^{(k-1)}[g]](x) = 0.$$

On the other hand, again following the style of the reasoning in the proof of Theorem 1, we can pull out the leading order terms, uniformly for $x \in E$ and $g \in B_1^+(E)$,

$$(50) \quad \begin{aligned} & \lim_{t \rightarrow \infty} \frac{1}{t^{2k-1}} \int_0^t \mathbf{T}_s \left[\beta \hat{\eta}_{t-s}^{(k-1)}[g] \right](x) ds \\ &= \lim_{t \rightarrow \infty} \int_0^1 \mathbf{T}_{ut} \left[\frac{\beta(\cdot)}{2} (1-u)^{2k-2} \mathcal{E} \cdot \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^{k-1} \binom{k}{k_i, k-k_i} \varphi(x_i) \varphi(x_j) \right. \right. \\ & \quad \left. \left. \times \frac{\mathbf{M}_{t(1-u)}^{(k_i)}[g](x_i)}{\varphi(x_i)(t(1-u))^{2k_i-1}} \frac{\mathbf{M}_{t(1-u)}^{(k-k_i)}[g](x_j)}{\varphi(x_j)(t(1-u))^{2k-2k_i-1}} \right] \right](x) du. \end{aligned}$$

It is again worth noting here that the choice of the polynomial growth in the form $t^{n(k)}$ also constrains the possible linear choices of $n(k)$ to $n(k) = 2k - 1$ if we are to respect $n(1) = 1$ and the correct distribution of the index across (50).

Identifying

$$F[g](x, u, t) = \frac{\beta(x)}{2\varphi(x)} (1-u)^{2k-2} \mathcal{E}_x \left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k_i=1}^{k-1} \binom{k}{k_i, k-k_i} \varphi(x_i) \varphi(x_j) \right. \\ \left. \times \frac{\mathbf{M}_{t(1-u)}^{(k_i)}[g](x_i)}{\varphi(x_i)(t(1-u))^{2k_i-1}} \frac{\mathbf{M}_{t(1-u)}^{(k-k_i)}[g](x_j)}{\varphi(x_j)(t(1-u))^{2k-2k_i-1}} \right],$$

our induction hypothesis allows us to conclude that $F[g](x, u) := \lim_{t \rightarrow \infty} F[g](x, u, t)$ exists and

$$\varphi(x)F[g](x, u) = (1 - u)^{2k-2} k! \frac{\beta(x)\mathbb{V}[\varphi](x)}{2^{k-1}} \langle g, \tilde{\varphi} \rangle^k \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \sum_{\ell=1}^{k-1} L_\ell L_{k-\ell}.$$

Thanks to our induction hypothesis, we can also easily verify (A.1) and (A.2). Theorem A.1 now gives us the required uniform (in $x \in E$ and $g \in B_1^+(E)$) limit

$$(51) \quad \lim_{t \rightarrow \infty} \frac{1}{t^{2k-1}} \int_0^t \mathbb{T}_s \left[\beta \hat{\eta}_{t-s}^{(k-1)}[g] \right] (x) ds = \frac{k! \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \langle g, \tilde{\varphi} \rangle^k}{2^{k-1}} L_k.$$

Putting (51) together with (49) we get the statement of Theorem 4. \square

2.5 Proofs for the non-critical cases

We now give an outline of the main steps in the proof of Theorem 1 for the sub and supercritical cases. As previously mentioned, the ideas used in this section will closely follow those presented in the previous section for the proof of the critical case and so we leave the details to the reader. We first note that the Perron Frobenius behaviour in (H1) ensures the base case for the induction argument, regardless of the value of λ . We thus turn to the inductive step, assuming the result holds for $k - 1$.

Proof of Theorem 2 (supercritical case). From the evolution equation (33), we have

$$(52) \quad \begin{aligned} & \lim_{t \rightarrow \infty} \frac{e^{-\lambda kt}}{\varphi(x)} \mathbb{T}_t^{(k)}[f](x) \\ &= \lim_{t \rightarrow \infty} \frac{e^{-\lambda kt}}{\varphi(x)} \int_0^t \mathbb{T}_s \left[\beta \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathbb{T}_{t-s}^{(k_j)}[f](x_j) \right] \right] (x) ds. \end{aligned}$$

It then follows that

$$(53) \quad \begin{aligned} & \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-k\lambda t} \mathbb{T}_t^{(k)}[f](x) - \langle f, \tilde{\varphi} \rangle^k L_k \right| \\ &= \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-k\lambda t} \int_0^t \mathbb{T}_s \left[\beta \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathbb{T}_{t-s}^{(k_j)}[f](x_j) \right] \right] (x) ds \right. \\ & \quad \left. - \int_0^t e^{-(k-1)\lambda s} ds \langle f, \tilde{\varphi} \rangle^k \lambda (k-1) L_k \right|. \end{aligned}$$

Noting that $\sum_{j=1}^N k_j = k$, we may again share the exponential term across the product in the right-hand side above as follows,

$$\begin{aligned} & \frac{e^{-\lambda kt}}{\varphi(x)} \int_0^t \mathbf{T}_s \left[\beta \mathcal{E} \left[\sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathbf{T}_{t-s}^{(k_j)} [f](x_j) \right] \right] (x) ds \\ &= t \int_0^1 e^{-\lambda(k-1)ut} \frac{e^{-\lambda ut}}{\varphi(x)} \mathbf{T}_{ut} \left[k! \beta \mathcal{E} \left(\sum_{[k_1, \dots, k_N]_k^2} \prod_{j=1}^N \varphi(x_j) \frac{e^{-\lambda k_j t(1-u)} \mathbf{T}_{t(1-u)}^{(k_j)} [f](x_j)}{k_j! \varphi(x_j)} \right) \right] (x) du. \end{aligned}$$

Combining this with to (53) and changing variables in the final integral of the latter, we have

$$\begin{aligned} & \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-k\lambda t} \mathbf{T}_t^{(k)} [f](x) - \langle f, \tilde{\varphi} \rangle^k L_k \right| \\ (54) \quad & \leq \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} t \left| \int_0^1 e^{-\lambda(k-1)ut} \left(\frac{e^{-\lambda ut}}{\varphi(x)} \mathbf{T}_{ut} [\varphi F(\cdot, u, t)] - \langle f, \tilde{\varphi} \rangle^k \lambda(k-1) L_k \right) du \right|, \end{aligned}$$

where we have defined

$$F[f](x, u, t) := k! \frac{\beta(x)}{\varphi(x)} \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^2} \prod_{j=1}^N \varphi(x_j) \frac{e^{-\lambda k_j t(1-u)} \mathbf{T}_{t(1-u)}^{(k_j)} [f](x_j)}{k_j! \varphi(x_j)} \right].$$

It is easy to see that, pointwise in $x \in E$ and $u \in [0, 1]$, using the induction hypothesis and (H2),

$$\begin{aligned} F[f](x, u) &:= \lim_{t \rightarrow \infty} F[f](x, u, t) \\ &= k! \frac{\beta(x)}{\varphi(x)} \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{\substack{j=1 \\ j: k_j > 0}}^N \varphi(x_j) L_{k_j} \right] \langle f, \tilde{\varphi} \rangle^k \end{aligned}$$

where we have again used the fact that the k_j s sum to k to extract the $\langle f, \tilde{\varphi} \rangle^k$ term. Similarly to the critical setting we can also verify using the induction hypothesis and (H2) that (A.1) and (A.2) hold.

This is sufficient to note that, by using a triangle inequality similar spirit to the one found in (A.4) and appealing to (13) of the assumption (H1), we have that

$$\sup_{x \in E, u \in [0, 1], f \in B_1^+(E), t \geq 0} \left| \frac{e^{-\lambda ut}}{\varphi(x)} \mathbf{T}_{ut} [\varphi F[f](\cdot, u, t)] - k! \langle f, \tilde{\varphi} \rangle^k L_k \right| < \infty.$$

This means that for t sufficiently large, we can control the modulus in the integral on the right-hand side of (54) by a global constant. The remainder of integral, yields a bound of $\varepsilon(1 - e^{-\lambda(k-1)t})/\lambda(k-1)$, which tends to zero as $t \rightarrow \infty$. \square

Proof of Theorem 3 (subcritical case). We now outline the subcritical case. First note that since we only compensate by $e^{-\lambda t}$, the term $\mathsf{T}_t[f^k](x)$ that appears in equation (33) does not vanish after the normalisation. Due to assumption (H1), we have

$$\lim_{t \rightarrow \infty} \varphi^{-1}(x) e^{-\lambda t} \mathsf{T}_t[f^k](x) = \langle f^k, \tilde{\varphi} \rangle.$$

Next we turn to the integral term in (33). Define $[k_1, \dots, k_N]_k^n$, for $2 \leq n \leq k$ to be the set of tuples (k_1, \dots, k_N) with exactly n positive terms and whose sum is equal to k . Similar calculations to those given above yield

$$\begin{aligned} & \frac{e^{-\lambda t}}{\varphi(x)} \int_0^t \mathsf{T}_s \left[\beta \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^n} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathsf{T}_{t-s}^{(k_j)}[f](x_j) \right] \right] (x) ds \\ (55) \quad & = t \sum_{n=2}^k \int_0^1 e^{\lambda(n-1)ut} \frac{e^{-\lambda t(1-u)}}{\varphi(x)} \mathsf{T}_{t(1-u)} \left[k! \beta \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^n} \prod_{j=1}^N \varphi(x_j) \frac{e^{-\lambda ut} \mathsf{T}_{ut}^{(k_j)}[f](x_j)}{k_j! \varphi(x_j)} \right] \right] (x) du \end{aligned}$$

Again, we leave the details to the reader but the idea is that the induction hypothesis will take care of the product of the lower order moments and the second part of (13) in assumption (H1) will then take care of the asymptotic behaviour semigroup $\mathsf{T}_{t(1-u)}$. The second part of (13) allows one to control the difference between this term and its limit. In a similar manner to the final step in the proof of Theorem 2, the difference of (33) and its limit can be reduced to the limit as $t \rightarrow \infty$ of $\varepsilon(1 - e^{-|\lambda|(n-1)t})/|\lambda|(n-1)$, which is bounded above by ε . \square

Proof of Theorem 5. For the case $k = 1$, we have

$$\begin{aligned} & \left| e^{-\lambda t} \varphi(x)^{-1} \int_0^t \mathsf{T}_s[g](x) ds - \frac{\langle g, \tilde{\varphi} \rangle}{\lambda} \right| \\ & = \left| e^{-\lambda t} t \int_0^1 e^{\lambda ut} (e^{-\lambda ut} \varphi(x)^{-1} \mathsf{T}_{ut}[g](x) - \langle g, \tilde{\varphi} \rangle) du - e^{-\lambda t} \frac{\langle g, \tilde{\varphi} \rangle}{\lambda} \right| \\ (56) \quad & \leq e^{-\lambda t} t \int_0^1 e^{\lambda ut} |e^{-\lambda ut} \varphi(x)^{-1} \mathsf{T}_{ut}[g](x) - \langle g, \tilde{\varphi} \rangle| du + e^{-\lambda t} \frac{\langle g, \tilde{\varphi} \rangle}{\lambda}. \end{aligned}$$

Thanks to (H1) and similar arguments to those used in the proof of Theorem 2, we may choose t sufficiently large such that the modulus in the integral is bounded above by $\varepsilon > 0$, uniformly in $g \in B_1^+(E)$ and $x \in E$. Then, the right-hand side of (56) is bounded above by $\varepsilon \lambda^{-1}(1 - e^{-\lambda t}) + e^{-\lambda t} \langle g, \tilde{\varphi} \rangle / \lambda$. Since ε can be taken arbitrarily small, this gives the desired result and also pins down the initial value $L_1 = 1/\lambda$.

Now assume the result holds for all $\ell \leq k-1$. Reflecting on proof of Theorem 2, we note that in this setting the starting point is almost identical except that the analogue of (52), which is derived from (47), is now the need to evaluate

$$\begin{aligned} & \lim_{t \rightarrow \infty} \frac{e^{-\lambda kt}}{\varphi(x)} \mathsf{M}_t^{(k)}[g](x) = \lim_{t \rightarrow \infty} \frac{e^{-\lambda kt}}{\varphi(x)} \int_0^t \mathsf{T}_s \left[\beta \mathcal{E}_x \left[\sum_{[k_1, \dots, k_N]_k^n} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \mathsf{M}_{t-s}^{(k_j)}[g](x_j) \right] \right] (x) ds \\ (57) \quad & - k \lim_{t \rightarrow \infty} \frac{e^{-\lambda kt}}{\varphi(x)} \int_0^t \mathsf{T}_s [g \mathsf{M}_{t-s}^{(k-1)}[g]](x) ds. \end{aligned}$$

The first term on the right-hand side of (57) can be handled in essentially the same way as in the proof of Theorem 2. The second term on the right-hand side of (57) can easily be dealt with along the lines that we are now familiar with from earlier proofs, using the induction hypothesis. In particular, its limit is zero. Hence combined with the first term on the right-hand side of (57), we recover the same recursion equation for L_k . \square

Proof of Theorem 6. The case $k = 1$ is relatively straightforward and, again in the interest of keeping things brief, we point the reader to the fact that, as $t \rightarrow \infty$, we have

$$(58) \quad \frac{1}{\varphi(x)} \mathbf{M}_t^{(1)}[g](x) = \int_0^t \frac{\mathbf{T}_s[g](x)}{\varphi(x)} ds = \int_0^t e^{\lambda s} e^{-\lambda s} \frac{\mathbf{T}_s[g](x)}{\varphi(x)} ds \sim t \langle g, \tilde{\varphi} \rangle \int_0^1 e^{\lambda ut} du \sim \frac{\langle g, \tilde{\varphi} \rangle}{|\lambda|}.$$

Now suppose the result holds for all $\ell \leq k - 1$. We again refer to (47), which means we are interested in handling a limit which is very similar to (57), now taking the form

$$(59) \quad \begin{aligned} & \frac{\mathbf{M}_t^{(k)}[g](x)}{\varphi(x)} \\ &= \frac{t}{\varphi(x)} \int_0^1 e^{\lambda ut} e^{-\lambda ut} \mathbf{T}_{ut} \left[\beta \mathcal{E} \left[\sum_{[k_1, \dots, k_N]_k^2} \binom{k}{k_1, \dots, k_N} \prod_{j=1}^N \varphi(x_j) \frac{\mathbf{M}_{t(1-u)}^{(k_j)}[g](x_j)}{\varphi(x_j)} \right] \right] (x) du \\ & \quad - k \frac{t}{\varphi(x)} \int_0^1 e^{\lambda ut} e^{-\lambda ut} \mathbf{T}_{ut} \left[g \varphi \frac{\mathbf{M}_{t(1-u)}^{(k-1)}[g]}{\varphi} \right] (x) du. \end{aligned}$$

Again skipping the details, we can quickly see from (59) the argument in (58), and the induction hypothesis gives us

$$(60) \quad \frac{\mathbf{M}_t^{(k)}[g](x)}{\varphi(x)} \sim k! \frac{\langle g, \tilde{\varphi} \rangle^k}{|\lambda|} \left\langle \beta \mathcal{E} \left[\sum_{[k_1, \dots, k_N]_k^2} \prod_{\substack{j=1 \\ j:k_j > 0}}^N \varphi(x_j) L_{k_j} \right], \tilde{\varphi} \right\rangle - k! \frac{\langle g \varphi, \tilde{\varphi} \rangle \langle g, \tilde{\varphi} \rangle^{k-1}}{|\lambda|} L_{k-1},$$

which gives us the required recursion for L_k . \square

3 Proofs for superprocesses

For the proof of Theorems 1, 2 and 3 in the setting of superprocesses we follow a similar approach. One difference is that we cannot work with the k -th moment as a product of an almost surely finite sum. As such the use of the Leibniz formula as in the previous section is no longer helpful. Instead, we use the Faà di Bruno formula (see Lemma A.1) to assist with multiple derivatives of the non-linear evolution equation (5).

3.1 Linear and non-linear semigroup equations

The evolution equation for the expectation semigroup $(\mathbf{T}_t, t \geq 0)$ is well known and satisfies

$$(61) \quad \mathbf{T}_t[f](x) = \mathbf{P}_t[f](x) + \int_0^t \mathbf{P}_s [\beta(\mathbf{m}[\mathbf{T}_{t-s}[f]] - 1) + b](x) ds,$$

for $t \geq 0$, $x \in \hat{E}$ and $f \in B^+(E)$, where, with a meaningful abuse of our branching Markov process notation, we now define

$$(62) \quad \begin{aligned} \mathfrak{m}[f](x) &= \int_{M_0(E)} \left[\gamma(x, \pi) \langle f, \pi \rangle + \int_0^\infty u \langle f, \pi \rangle n(x, \pi, du) \right] G(x, d\pi) \\ &= \gamma(x, f) + \int_{M(E)^\circ} \langle f, \nu \rangle \Gamma(x, \nu). \end{aligned}$$

See for example equation (3.24) of [4].

In the spirit of Lemma 1 we can give a second representation of $\mathsf{T}_t[f]$ in terms of an auxiliary process, the so called many-to-one formula. To this end, if, as before, we work with the process (ξ, \mathbf{P}) to represent the Markov process associated to the semigroup $(\mathsf{P}_t, t \geq 0)$, then, although we have redefined the quantity $\mathfrak{m}[f](x)$, we can still meaningfully work with the process $(\hat{\xi}, \hat{\mathbf{P}})$ as defined just before Lemma 1.

Lemma 3. *Let $\vartheta(x) = B(x) + b(x) = \beta(x)(\mathfrak{m}[1](x) - 1) + b(x)$, then, for $t \geq 0$ and $f \in B^+(E)$,*

$$(63) \quad \mathsf{T}_t[f](x) = \hat{\mathbf{E}}_x \left[\exp \left(\int_0^t \vartheta(\hat{\xi}_s) ds \right) f(\hat{\xi}_t) \right].$$

As with Lemma 1, the proof is classical, requiring only that we take the right-hand side of (63) and condition on the first extra jump of $(\hat{\xi}, \hat{\mathbf{P}})$ to show that it also solves (62). It is a straightforward application of Grönwall's inequality to show that (62) has a unique solution and hence (61) holds. The reader will note that because we have separated out the local and non-local branching mechanisms of the superprocess, the deliberate repeat definition of $\mathfrak{m}[f]$ for superprocesses is only the analogue of its counter part for branching Markov processes in the sense of non-local activity. The mean local branching rate has otherwise been singled out as the term b .

Similarly to the branching Markov process setting, let us re-write an extended version of the non-linear semigroup evolution $(\mathsf{V}_t, t \geq 0)$, defined in (6), i.e. the natural analogue of (27), in terms of the linear semigroup $(\mathsf{T}_t, t \geq 0)$. To this end, define

$$\mathsf{V}_t[f, g](x) = \mathbb{E}_x \left[e^{-\langle f, X_t \rangle - \int_0^t \langle g, X_s \rangle ds} \right],$$

Analogously to Theorem 2 we have the following result.

Lemma 4. *For all $f, g \in B^+(E)$, $x \in E$ and $t \geq 0$, the non-linear semigroup $\mathsf{V}_t[f, g](x)$ satisfies*

$$(64) \quad \mathsf{V}_t[f, g](x) = \mathsf{T}_t[f](x) - \int_0^t \mathsf{T}_s [\mathsf{J}[\mathsf{V}_{t-s}[f, g]] - g\mathsf{V}_{t-s}[f, g]](x) ds,$$

where, for $h \in B^+(E)$ and $x \in E$,

$$\mathsf{J}[h](x) = \psi(x, h(x)) + \phi(x, h) + \beta(x)(\mathfrak{m}[h](x) - h(x)) + b(x)h(x).$$

The proof is essentially the same as the proof of Lemma 2 and hence we leave the details to the reader.

3.2 Evolution equations for the k -th moment of a superprocesses

Recall that we defined $\mathbf{T}_t^{(k)}[f](x) := \mathbb{E}_{\delta_x}[\langle f, X_t \rangle^k]$, $t \geq 0$, $f \in B^+(E)$, $k \geq 1$. As with the setting of branching Markov processes, we want to establish an evolution equation for $(\mathbf{T}_t^{(k)}, t \geq 0)$, from which we can establish the desired asymptotics. To this end, let us introduce the following notation.

For $x \in E$, $k \geq 2$ and $t \geq 0$, define

$$(65) \quad R_k(x, t) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} (-1)^{m_1 + \dots + m_{k-1} - 1} (m_1 + \dots + m_{k-1} - 1)! \prod_{j=1}^{k-1} \left(\frac{(-1)^j \mathbf{T}_t^{(j)}[f](x)}{j!} \right)^{m_j},$$

and

$$(66) \quad K_k(x, t) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \psi^{(m_1 + \dots + m_{k-1})}(x, 0+) \prod_{j=1}^{k-1} \left(\frac{(-1)^{j+1} \mathbf{T}_t^{(j)}[f](x) - R_j(x, t)}{j!} \right)^{m_j},$$

and finally

$$(67) \quad S_k(x, t) = \int_{M(E)^\circ} \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} (-1)^{m_1 + \dots + m_{k-1}} \prod_{j=1}^{k-1} \left(\frac{\langle (-1)^{j+1} \mathbf{T}_t^{(j)}[f] - R_j(\cdot, t), \nu \rangle}{j!} \right)^{m_j} \Gamma(x, d\nu),$$

and the sums run over the set of non-negative integers $\{m_1, \dots, m_{k-1}\}$ such that $m_1 + 2m_2 + \dots + (k-1)m_{k-1} = k$.

Theorem 7. Fix $k \geq 2$. Suppose that (H1) and (H2) hold, with the additional assumption that

$$(68) \quad \sup_{x \in E, s \leq t} \mathbf{T}_s^{(\ell)}[f](x) < \infty, \quad \ell \leq k-1, f \in B^+(E), t \geq 0.$$

Then,

$$(69) \quad \mathbf{T}_t^{(k)}[f](x) = (-1)^{k+1} R_k(x, t) + (-1)^k \int_0^t \mathbf{T}_s[U_k(\cdot, t-s)] ds,$$

where

$$(70) \quad U_k(x, t) = K_k(x, t) + \beta(x) S_k(x, t).$$

Proof. For the proof of this result, recall the definition (5) and let

$$\mathbf{v}_t^{(k)}[f](x) := \frac{\partial^k}{\partial \theta^k} \mathbf{V}_t[\theta f, 0](x) \Big|_{\theta=0}, \quad t \geq 0, f \in B^+(E), k \geq 1$$

as well as

$$\mathbf{e}_t[f](x) := \mathbb{E}_{\delta_x} [e^{-\langle f, X_t \rangle}], \quad t \geq 0, f \in B^+(E).$$

In that case, $\mathbf{V}_t[\theta f](x) = -\log \mathbf{e}_t[\theta f](x)$ and $\mathbf{e}_t[0](x) = 1$, so that

$$\mathbf{e}_t^{(k)}[\theta f](x) := \frac{\partial^k}{\partial \theta^k} \mathbf{e}_t[\theta f](x) = (-1)^k \mathbb{E}_{\delta_x} [\langle f, X_t \rangle^k e^{-\theta \langle f, X_t \rangle}]$$

and

$$(71) \quad \mathbf{e}_t^{(k)}[\theta f](x)|_{\theta=0} = (-1)^k \mathbf{T}_t^{(k)}[f](x).$$

Next we can use Faà di Bruno's Lemma A.1 to get

$$\begin{aligned} & \mathbf{v}_t^{(k)}[f](x) \\ &= \frac{\partial^k}{\partial \theta^k} -\log \mathbf{e}_t[\theta f](x) \Big|_{\theta=0} \\ &= - \sum_{\{m_1, \dots, m_k\}_k} \frac{k!}{m_1! \dots m_k!} \frac{(-1)^{m_1 + \dots + m_k - 1} (m_1 + \dots + m_k - 1)!}{\mathbf{e}_t[\theta f](x)^{m_1 + \dots + m_k}} \prod_{j=1}^k \left(\frac{\mathbf{e}_t^{(j)}[\theta f]}{j!} \right)^{m_j} \Big|_{\theta=0} \\ &= - \sum_{\{m_1, \dots, m_k\}_k} \frac{k!}{m_1! \dots m_k!} (-1)^{m_1 + \dots + m_k - 1} (m_1 + \dots + m_k - 1)! \prod_{j=1}^k \left(\frac{(-1)^j \mathbf{T}_t^{(j)}[f](x)}{j!} \right)^{m_j}, \end{aligned}$$

where the sum runs over the set of non-negative integers $\{m_1, \dots, m_k\}_k$ such that

$$m_1 + 2m_2 + \dots + km_k = k.$$

Note that $m_k > 0$ if and only if $m_k = 1$ and $m_1 = m_2 = \dots = m_{k-1} = 0$, so the k -th moment term $\mathbf{T}_t^{(k)}[f]$ appears only once and with a factor $(-1)^{k+1}$, that is,

$$(72) \quad \mathbf{v}_t^{(k)}[f](x) = (-1)^{k+1} \mathbf{T}_t^{(k)}[f](x) - R_k(x, t),$$

where all the terms in $R_k(x, t)$ are products of two or more lower order moments.

Now, we differentiate the evolution equation (64) k times at $\theta = 0$, momentarily not worrying about passing derivatives through integrals, to get that

$$\mathbf{v}_t^{(k)}[f](x) = - \int_0^t \mathbf{T}_s \left[\frac{\partial^k}{\partial \theta^k} \left(\psi(\cdot, \mathbf{V}_{t-s}[\theta f, 0](\cdot)) + \phi(\cdot, \mathbf{V}_{t-s}[\theta f, 0]) + \mathbf{F}[\mathbf{V}_{t-s}[\theta f, 0]] \right) \Big|_{\theta=0} \right] (x) ds,$$

where

$$\mathbf{F}[g](x) = \beta(x)(m[g] - g) + b(x)g, \quad x \in E, g \in B^+(E).$$

For the k th derivative of $\psi(x, \mathbf{V}_t[\theta f, 0](x))$ at $\theta = 0$ we again use Faà di Bruno, Lemma A.1, to get

$$\begin{aligned} & \left. \frac{\partial^k}{\partial \theta^k} \psi(x, \mathbf{V}_t[\theta f, 0](x)) \right|_{\theta=0} \\ &= \sum_{\{m_1, \dots, m_k\}_k} \frac{k!}{m_1! \dots m_k!} \psi^{(m_1 + \dots + m_k)}(x, \mathbf{V}_t[\theta f, 0]) \prod_{j=1}^k \left(\frac{\frac{\partial^j}{\partial \theta^j} \mathbf{V}_t[\theta f, 0](x)}{j!} \right)^{m_j} \Big|_{\theta=0} \\ &= \sum_{\{m_1, \dots, m_k\}_k} \frac{k!}{m_1! \dots m_k!} \psi^{(m_1 + \dots + m_k)}(x, 0+) \prod_{j=1}^k \left(\frac{\mathbf{v}_t^{(j)}[f](x)}{j!} \right)^{m_j} \\ &= -b(x) \mathbf{v}_t^{(k)}[f](x) + K_k(x, t), \end{aligned}$$

where the last equality holds because $m_k = 1$ if and only if $m_1 = \dots = m_{k-1} = 0$ and $\psi'(x, 0+) = -b(x)$. Similarly, for the the k th derivative of the remaining terms recalling (8), (9) and (62),

$$\begin{aligned} & \frac{\partial^k}{\partial \theta^k} \left(\phi(x, \mathbf{V}_t[\theta f, 0]) + \mathbf{F}[\mathbf{V}_t[\theta f, 0]] \right) \\ &= b(x) \frac{\partial^k}{\partial \theta^k} \mathbf{V}_t[\theta f, 0] \\ & \quad - \beta(x) \int_{M_0(E)} \int_0^\infty \frac{\partial^k}{\partial \theta^k} \left(1 - e^{-u \langle \mathbf{V}_t[\theta f, 0], \pi \rangle} - u \langle \mathbf{V}_t[\theta f, 0], \pi \rangle \right) n(x, \pi, du) G(x, d\pi), \end{aligned}$$

and using Lemma A.1 we have

$$\begin{aligned} & \frac{\partial^k}{\partial \theta^k} \left(1 - e^{-u \langle \mathbf{V}_t[\theta f, 0], \pi \rangle} - u \langle \mathbf{V}_t[\theta f, 0], \pi \rangle \right) \\ &= \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_k!} (-1)^{m_1 + \dots + m_k + 1} e^{-u \langle \mathbf{V}_t[\theta f, 0], \pi \rangle} \prod_{j=1}^k \left(\frac{u \langle \frac{\partial^j}{\partial \theta^j} \mathbf{V}_t[\theta f, 0], \pi \rangle}{j!} \right)^{m_j} \\ & \quad + \left(e^{-u \langle \mathbf{V}_t[\theta f, 0], \pi \rangle} - 1 \right) u \left\langle \frac{\partial^k}{\partial \theta^k} \mathbf{V}_t[\theta f, 0], \pi \right\rangle \end{aligned}$$

where, in the final equality, we have singled out the case that $m_k = 1$ and $m_1 = \dots = m_{k-1} = 0$ in the Faà di Bruno formula. and then, using the definition of $m[f](x)$ in (62) and the same observation as above about the m_j 's, we get

$$(73) \quad \left. \frac{\partial^k}{\partial \theta^k} \left(\phi(x, \mathbf{V}_t[\theta f, 0]) + \mathbf{F}[\mathbf{V}_t[\theta f, 0]] \right) \right|_{\theta=0} = b(x) \mathbf{v}_t^{(k)}[f](x) + \beta(x) S_k(x, t).$$

Putting the pieces together, we get

$$(74) \quad \mathbf{v}_t^{(k)}[f](x) = - \int_0^t \mathbf{T}_s [U_k(\cdot, t-s)](x) ds.$$

Combining this with equation (72) we get that

$$(-1)^{k+1} \mathbf{T}_t^{(k)} [f](x) = R_k(x, t) - \int_0^t \mathbf{T}_s [U_k(\cdot, t-s)](x) ds,$$

which is the desired result.

There is one final matter we must attend to, which is the ability to move derivatives through integrals. In this setting, this is easier to deal with thanks to the the assumption (68), (H2) and the Lévy-Khintchine-type formulae for ψ and ϕ . \square

3.3 Completing the proof of Theorem 1: critical case

We will prove Theorem 1 for superprocesses using induction, similarly to the setting of branching Markov processes. The case $k = 1$ follows from assumption (H1).

Now assume that the statement of Theorem 1 holds in the superprocess setting for all $\ell \leq k$. Our aim is to prove that the result holds for $k + 1$. Using Theorem 7 and a change of variables, we have that

$$(75) \quad \frac{1}{\varphi(x)t^k} \mathbf{T}_t^{(k+1)} [f](x) = \frac{(-1)^k}{\varphi(x)t^k} R_{k+1}(x, t) + \frac{(-1)^{k+1}}{\varphi(x)t^{k-1}} \int_0^1 \mathbf{T}_{st} [U_{k+1}(\cdot, t(1-s))](x) ds,$$

where R and U were defined in equations (65) and (70), respectively. We will prove first that, for each $x \in E$,

$$(76) \quad \begin{aligned} & \lim_{t \rightarrow \infty} \frac{1}{\varphi(x)t^k} \mathbf{T}_t^{(k+1)} [f](x) \\ &= \lim_{t \rightarrow \infty} \frac{1}{2\varphi(x)t^{k-1}} \int_0^1 \mathbf{T}_{st} \left[K_{k+1}^{(2)}(\cdot, t(1-s)) + \beta(\cdot) S_{k+1}^{(2)}(\cdot, t(1-s)) \right] (x) ds, \end{aligned}$$

where

$$(77) \quad K_{k+1}^{(2)}(x, t) := \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1!k_2!} \psi''(x, 0+) \mathbf{T}_t^{(k_1)} [f](x) \mathbf{T}_t^{(k_2)} [f](x)$$

and

$$(78) \quad S_{k+1}^{(2)}(x, t) = \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1!k_2!} \langle \mathbf{T}_t^{(k_1)} [f], \nu \rangle \langle \mathbf{T}_t^{(k_2)} [f], \nu \rangle \Gamma(x, d\nu),$$

such that $\{k_1, k_2\}^+$ is defined to be the set of positive integers k_1, k_2 such that $k_1 + k_2 = k + 1$.

To this end, writing $c(m_1, \dots, m_k)$ for the constants preceding the product summands in (65), observe that

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t^k} R_{k+1}(x, t) &= \lim_{t \rightarrow \infty} \frac{(k+1)!}{t^k} \sum_{\{m_1, \dots, m_k\}_{k+1}} c(m_1, \dots, m_k) \prod_{j=1}^k \left(\frac{(-1)^j \mathbf{T}_t^{(j)}[f](x)}{j!} \right)^{m_j} \\ &= (-1)^k (k+1)! \lim_{t \rightarrow \infty} \sum_{\{m_1, \dots, m_k\}_{k+1}} \frac{c(m_1, \dots, m_k)}{t^{m_1 + \dots + m_k - 1}} \prod_{j=1}^k \left(\frac{1}{j!} \frac{\mathbf{T}_t^{(j)}[f](x)}{t^{j-1}} \right)^{m_j} \\ &= 0, \end{aligned}$$

where the final equality is due to the induction hypothesis and the fact that $m_1 + \dots + m_k > 1$, which follows from the fact that $m_1 + 2m_2 + \dots + km_k = k + 1$. Note, moreover that the induction hypothesis ensures that the limit is uniform in $x \in E$ and, in fact, that

$$(79) \quad \sup_{t \geq 0, x \in E} \frac{1}{t^{\ell-1}} R_\ell(x, t) < \infty \text{ and } \limsup_{t \rightarrow \infty} \sup_{x \in E} \frac{1}{t^{\ell-1}} R_\ell(x, t) = 0 \quad \ell = 1, \dots, k+1.$$

We now return to (75), to deal with the term involving U_{k+1} , which we recall is a linear combination of K_{k+1} and S_{k+1} , which were defined in (65) and (67), respectively. Note that if any of the summands in either K_{k+1} or S_{k+1} have more than two of the m_j positive, the limit of that summand, when renormalised by $1/t^{k-1}$, will be zero. In essence, the argument here is analogous to those that led to (41) in the branching Markov process setting. This implies that the only terms in the sums of (65) and (67) that remain in the limit of (75) are those for which $m_{k_1} = m_{k_2} = 1$ and $m_j = 0$ for all $j \neq k_1, k_2$, with $k_1 < k_2$ such that $k_1 + k_2 = k + 1$, and if $k + 1$ is even, the terms in which $m_{(k+1)/2} = 2$ and $m_j = 0$ for all $j \neq (k+1)/2$.

Let us now convert all of the above heuristics into rigorous computation. We write

$$(80) \quad F[f](x, s, t) := \frac{1}{\varphi(x)t^{k-1}} \left(K_{k+1}^{(3+)}(x, t(1-s)) + \beta(x) S_{k+1}^{(3+)}(x, t(1-s)) \right),$$

where $K_{k+1}^{(3+)}$ and $S_{k+1}^{(3+)}$ contain the terms in K_{k+1} and S_{k+1} , respectively, for which the sum $m_1 + \dots + m_k$ is greater than or equal to 3. We will prove that $\lim_{t \rightarrow \infty} F(x, s, t) = 0$ and that (A.1) and (A.2) hold.

Due to (15) and boundness of φ , dominated convergence implies that

$$\begin{aligned} &\lim_{t \rightarrow \infty} \frac{1}{\varphi(x)t^{k-1}} K_{k+1}^{(3+)}(x, t(1-s)) \\ &= \frac{(k+1)!}{\varphi(x)} \sum_{\{m_1, \dots, m_k\}_{k+1}^3} \frac{\psi^{(m_1 + \dots + m_k)}(x, 0+)}{m_1! \dots m_k!} \\ &\quad \lim_{t \rightarrow \infty} \frac{1}{t^{m_1 + \dots + m_k - 2}} \prod_{j=1}^k \left(\frac{(-1)^{j+1} \mathbf{T}_{t(1-s)}^{(j)}[f](x) - R_j(x, t(1-s))}{j! t^{j-1}} \right)^{m_j}, \end{aligned}$$

where the set $\{m_1, \dots, m_k\}_{k+1}^3$ is the subset of $\{m_1, \dots, m_k\}_{k+1}$ for which $m_1 + \dots + m_k \geq 3$. Using the induction hypothesis and (79), we get that the right-hand side above is zero. The same arguments also imply that the limit of $S_{k+1}^{(3+)}$ is zero. Thus $F[f](x, s) := \lim_{t \rightarrow \infty} F[f](x, s, t) = 0$. The condition (A.1) trivially holds. For (A.2), the required uniformity follows from the induction hypothesis and (79).

Using Theorem A.1 in the Appendix, we conclude that

$$\limsup_{t \rightarrow \infty} \sup_{x \in E} \left| \frac{1}{\varphi(x)t^{k-1}} \int_0^1 \mathbf{T}_{ts} [\varphi F[f](\cdot, s, t)](x) ds \right| = 0.$$

Let us now define $\{k_1 < k_2\}$ to be the elements in K_{k+1} for which $m_{k_1} = m_{k_2} = 1$ with $k_1 < k_2$ such that $k_1 + k_2 = k + 1$ and $m_j = 0$ for all other indices and, in the case where $k + 1$ is even, $m_{(k+1)/2} = 2$ and $m_j = 0$ for all $j \neq (k + 1)/2$. Restricting the sum to this set in K_{k+1} we get the following expression

$$\begin{aligned} K_{k+1}^{(2)}(x, t) &= \sum_{\{k_1 < k_2\}} \frac{(k+1)!}{k_1!k_2!} \psi''(x, 0+) (-1)^{k+1} \mathbf{T}_t^{(k_1)}[f](x) \mathbf{T}_t^{(k_2)}[f](x) \\ &\quad + \mathbf{1}_{(k+1 \text{ is even})} \frac{1}{2} \binom{k+1}{k/2} \psi''(x, 0+) (-1)^{k+1} \left(\mathbf{T}_t^{(k/2)}[f](x) \right)^2 \\ &= \frac{(-1)^{k+1}}{2} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1!k_2!} \psi''(x, 0+) \mathbf{T}_t^{(k_1)}[f](x) \mathbf{T}_t^{(k_2)}[f](x), \end{aligned}$$

where we recall $\{k_1, k_2\}^+$ is the set of positive integers k_1, k_2 such that $k_1 + k_2 = k + 1$. Similarly, we obtain the following expression for S_{k+1} :

$$\begin{aligned} K_{k+1}^{(2)}(x, t) &= \int_{M(E)^\circ} \sum_{k_1 < k_2} \frac{(k+1)!}{k_1!k_2!} (-1)^{k+1} \langle \mathbf{T}_t^{(k_1)}[f], \nu \rangle \langle \mathbf{T}_t^{(k_2)}[f], \nu \rangle \Gamma(x, d\nu) \\ &\quad + \mathbf{1}_{(k+1 \text{ is even})} \int_{M(E)^\circ} \binom{k+1}{k/2} \frac{(-1)^{k+1}}{2} \langle \mathbf{T}_t^{((k+1)/2)}[f], \nu \rangle^2 \Gamma(x, d\nu) \\ &= \frac{(-1)^{k+1}}{2} \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1!k_2!} \langle \mathbf{T}_t^{(k_1)}[f], \nu \rangle \langle \mathbf{T}_t^{(k_2)}[f], \nu \rangle \Gamma(x, d\nu). \end{aligned}$$

Combining this with (75), we obtain (76). To conclude the proof define

$$(81) \quad F[f](x, s, t) := \frac{1}{2\varphi(x)t^{k-1}} \left(K_{k+1}^{(2)}(x, t(1-s)) + \beta(x) S_2^{(2)}(x, t(1-s)) \right).$$

Due to (15) and the induction hypothesis,

$$\begin{aligned}
& \lim_{t \rightarrow \infty} \frac{1}{2\varphi(x)t^{k-1}} K_{k+1}^{(2)}(x, t(1-s)) \\
&= \frac{(1-s)^{k-1}}{2\varphi(x)} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1!k_2!} \psi''(x, 0+) \lim_{t \rightarrow \infty} \frac{\mathbf{T}_{t(1-s)}^{(k_1)} f(x)}{(t(1-s))^{k_1-1}} \frac{\mathbf{T}_{t(1-s)}^{(k_2)} f(x)}{(t(1-s))^{k_2-1}} \\
&= (1-s)^{k-1} \varphi(x) \sum_{\{k_1, k_2\}^+} (k+1)! 2^{-k} \psi''(x, 0+) \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \\
&= k(1-s)^{k-1} \varphi(x) (k+1)! 2^{-k} \psi''(x, 0+) \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1},
\end{aligned}$$

where the last equality holds because the total number of ways of splitting one set of size $k+1$ into two non empty sets is equal to k . To obtain the limit for $S_{k+1}^{(2)}$, we use (15), the induction hypothesis, dominated convergence and linearity to obtain

$$\begin{aligned}
& \lim_{t \rightarrow \infty} \frac{S_{k+1}^{(2)}(x, t(1-s))}{2\varphi(x)t^{k-1}} \\
&= \frac{(1-s)^{k-1}}{2\varphi(x)} \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} \frac{(k+1)!}{k_1!k_2!} \lim_{t \rightarrow \infty} \frac{\langle \mathbf{T}_{t(1-s)}^{(k_1)} [f], \nu \rangle \langle \mathbf{T}_{t(1-s)}^{(k_2)} [f], \nu \rangle}{(t(1-s))^{k_1-1} (t(1-s))^{k_2-1}} \Gamma(x, d\nu) \\
&= \frac{(1-s)^{k-1}}{2^k \varphi(x)} \int_{M(E)^\circ} \sum_{\{k_1, k_2\}^+} (k+1)! \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \langle \varphi, \nu \rangle^2 \Gamma(x, d\nu) \\
&= \frac{k(1-s)^{k-1}}{2^k \varphi(x)} (k+1)! \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \int_{M(E)^\circ} \langle \varphi, \nu \rangle^2 \Gamma(x, d\nu).
\end{aligned}$$

Combining these two limits, we get that

$$F[f](x, s) := \lim_{t \rightarrow \infty} F[f](x, s, t) = \frac{k(1-s)^{k-1} (k+1)!}{\varphi(x) 2^k} \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \mathbb{V}[\varphi](x).$$

In order to complete the proof, we will use Theorem A.1 to deal with (76). By now the reader will be familiar with the arguments required to verify assumptions (A.1) and (A.2) and thus, we exclude the details. Hence, it follows that

$$\begin{aligned}
\lim_{t \rightarrow \infty} \frac{1}{\varphi(x)t^k} \mathbf{T}^{(k+1)} [f](x) &= \frac{(k+1)!}{2^k} \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^{k-1} \int_0^1 k(1-s)^{k-1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle ds \\
&= \frac{(k+1)!}{2^k} \langle f, \tilde{\varphi} \rangle^{k+1} \langle \mathbb{V}[\varphi], \tilde{\varphi} \rangle^k,
\end{aligned}$$

where the limit is uniform in $x \in E$. Moreover, $\sup_{t \geq 0, x \in E} \mathbf{T}^{(k+1)} [f](x) / \varphi(x)t^k < \infty$.

3.4 Proofs for the non-critical cases

In this section we present the main ideas behind the proof of Theorems 2 and 3. The methods follow a similar reasoning to the critical case and the details are left to the reader. The base

case is given by the Perron Frobenius behaviour in (H1) for both sub and supercritical cases. Thus, we assume the result for $k - 1$ and proceed to give the outline of the inductive step of the argument.

Proof of Theorem 2 (supercritical case). The main difference now compared to the critical case is that all the terms in $R_k(x, t)$ will survive after the normalization $e^{-\lambda kt}$ as the exponential term shares across the product. From the evolution equation (69) and the definition of L_k we have that

$$\begin{aligned} & |\varphi(x)^{-1} e^{-\lambda kt} \mathbf{T}_t^{(k)} [f](x) - k! \langle f, \tilde{\varphi} \rangle^k L_k(x)| \\ & \leq \left| \varphi(x)^{-1} e^{-\lambda kt} (-1)^{k+1} R_k(x, t) \right. \\ & \quad \left. - \langle f, \tilde{\varphi} \rangle^k \sum_{\{m_1, \dots, m_{k-1}\}} \frac{k!}{m_1! \dots m_{k-1}!} (m_1 + \dots + m_{k-1} - 1)! \varphi(x)^{m_1 + \dots + m_{k-1} - 1} \prod_{j=1}^{k-1} (-L_j(x))^{m_j} \right. \\ & \quad \left. + \left| \varphi(x)^{-1} e^{-\lambda kt} (-1)^k \int_0^t T_s [U_k(\cdot, t - s)](x) ds - \frac{k! \langle f, \tilde{\varphi} \rangle^k}{\lambda(k-1)} \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right| \right|. \end{aligned}$$

The first terms in the right hand side goes to zero uniformly since

$$\begin{aligned} \frac{e^{-\lambda kt}}{\varphi(x)} (-1)^{k+1} R_k(x, t) &= \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} (-1)^{m_1 + \dots + m_{k-1}} \\ & \quad \times (m_1 + \dots + m_{k-1} - 1)! \prod_{j=1}^{k-1} \left(\frac{e^{-\lambda j t} \mathbf{T}_t^{(j)} [f](x)}{\varphi(x) j!} \right)^{m_j} \varphi(x)^{m_1 + \dots + m_{k-1} - 1}, \end{aligned}$$

and the induction hypothesis implies that

$$\begin{aligned} (82) \quad \lim_{t \rightarrow \infty} \frac{e^{-\lambda kt}}{\varphi(x)} (-1)^{k+1} R_k(x, t) &= \langle f, \tilde{\varphi} \rangle^k \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \\ & \quad \times (m_1 + \dots + m_{k-1} - 1)! \varphi(x)^{m_1 + \dots + m_{k-1} - 1} \prod_{j=1}^{k-1} (-L_j(x))^{m_j}. \end{aligned}$$

For the second term, define for $k \geq 2$

$$I_k(x, t) := \int_0^t T_s [U_k(\cdot, t - s)](x) ds = (-1)^k \mathbf{T}_t^{(k)} [f](x) + R_k(x, t).$$

We will use induction to prove that for $k \geq 2$

$$(83) \quad \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \frac{e^{-\lambda kt}}{\varphi(x)} (-1)^k I_k(x, t) - \frac{k! \langle f, \tilde{\varphi} \rangle^k}{\lambda(k-1)} \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right| = 0,$$

which will complete the proof of the theorem.

First notice that for any $k \geq 2$ due to a change of variable, we have that

$$\frac{e^{-\lambda kt}}{\varphi(x)} (-1)^k I_k(x, t) = t \int_0^1 e^{-\lambda(k-1)ut} \frac{e^{-\lambda ut}}{\varphi(x)} \mathbf{T}_{ut} [(-1)^k e^{-\lambda kt(1-u)} U_k(\cdot, t(1-u))] (x) du,$$

where

$$(-1)^k e^{-\lambda kt} U_k(x, t) = (-1)^k e^{-\lambda kt} K_k(x, t) + \beta(x) (-1)^k e^{-\lambda kt} S_k(x, t).$$

Recalling the definitions of K_k and S_k given in (66) and (67) respectively, and using the fact that $I_j(x, t) = -((-1)^{j+1} \mathbf{T}_t[f](x) - R_j(x, t))$ for $j = 2, \dots, k-1$, after sharing the exponential term across the products, we obtain

$$(-1)^k e^{-\lambda kt} K_k(x, t) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \psi^{(m_1 + \dots + m_{k-1})}(x, 0+) (-e^{-\lambda t} \mathbf{T}_t[f](x))^{m_1} \prod_{j=2}^{k-1} \left(-\frac{1}{j!} (-1)^j e^{-\lambda jt} I_j(x, t) \right)^{m_j}$$

and

$$(-1)^k e^{-\lambda kt} S_k(x, t) = \int_{M(E)^\circ} \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} (-1)^{m_1 + \dots + m_{k-1}} \langle -e^{-\lambda t} \mathbf{T}_t[f], \nu \rangle^{m_1} \prod_{j=2}^{k-1} \frac{1}{j!} \langle -(-1)^j e^{-\lambda jt} I_j(\cdot, t), \nu \rangle^{m_j} \Gamma(x, d\nu).$$

From these expressions and the definition of \mathbb{V}_2 , the case $k = 2$ follows easily. Now we assume (83) holds for $\ell = 1, \dots, k-1$, then similarly to the Markov branching process case it follows that

$$(84) \quad \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \frac{e^{-\lambda kt}}{\varphi(x)} (-1)^k I_k - \frac{k! \langle f, \tilde{\varphi} \rangle^k}{\lambda(k-1)} \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right| \leq \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} t \left| \int_0^1 e^{-\lambda(k-1)ut} \left(\frac{e^{-\lambda ut}}{\varphi(x)} \mathbf{T}_{ut} [\varphi F(\cdot, u, t)] - k! \langle f, \tilde{\varphi} \rangle^k \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right) du \right|,$$

where we have defined

$$\begin{aligned} F[f](x, u, t) &= \varphi(x)^{-1} \left((-1)^k e^{-\lambda kt(1-u)} K_k(x, t(1-u)) + \beta(x) (-1)^k e^{-\lambda kt(1-u)} S_k(x, t(1-u)) \right) \\ &= \frac{1}{\varphi(x)} \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \\ &\times \left[\psi^{(m_1 + \dots + m_{k-1})} (-\varphi)^{m_1 + \dots + m_{k-1}} \left(\frac{e^{-\lambda t(1-u)} \mathbf{T}_{t(1-u)}[f](x)}{\varphi(x)} \right)^{m_1} \prod_{j=2}^{k-1} \left(\frac{(-1)^j e^{-\lambda jt(1-u)} I_j(x, t(1-u))}{\varphi(x) j!} \right)^{m_j} \right. \\ &\left. + \beta(x) \int_{M(E)^\circ} \left\langle \frac{e^{\lambda t(1-u)} \mathbf{T}_{t(1-u)}[f]}{\varphi} \varphi, \nu \right\rangle^{m_1} \prod_{j=2}^{k-1} \frac{1}{j!} \left\langle \frac{(-1)^j e^{-\lambda jt(1-u)} I_j(\cdot, t(1-u))}{\varphi} \varphi, \nu \right\rangle^{m_j} \Gamma(x, d\nu) \right]. \end{aligned}$$

It is easy to see that, pointwise in $x \in E$ and for $u \in (0, 1)$, using the induction hypothesis for I_k and the assumed Perron Frobenius behaviour (H1) for $\mathbf{T}_{t(1-u)}[f](x)$ we have

$$\begin{aligned}
F[f](x, u) &:= \lim_{t \rightarrow \infty} F[f](x, u, t) \\
&= \frac{1}{\varphi(x)} \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \langle f, \tilde{\varphi} \rangle^{m_1} \prod_{j=2}^{k-1} \left(\frac{\langle f, \tilde{\varphi} \rangle^j \langle \mathbb{V}_j[\varphi], \tilde{\varphi} \rangle}{\lambda(j-1)} \right)^{m_j} \\
&\quad \times \left[\psi^{(m_1 + \dots + m_{k-1})}(x, 0+) (-\varphi(x))^{m_1 + \dots + m_{k-1}} + \beta(x) \int_{M(E)^\circ} \langle \varphi, \nu \rangle^{m_1 + \dots + m_{k-1}} \Gamma(x, d\nu) \right] \\
&= \frac{k!}{\varphi(x)} \langle f, \tilde{\varphi} \rangle^k \mathbb{V}_k[\varphi](x).
\end{aligned}$$

We again verify that the conditions (A.1) and (A.2) hold using the induction hypothesis and (H2). To complete the proof of (83), we again proceed along the same lines as in the branching Markov processes setting. Similar arguments to those given in the proof of Theorem A.1

$$\sup_{\substack{x \in E, u \in [0, 1], \\ f \in B_1^+(E), t \geq 0}} \left| \frac{e^{-\lambda u t}}{\varphi(x)} \mathbf{T}_{ut} [\varphi F[f](x, u, t)] - k! \langle f, \tilde{\varphi} \rangle^k \mathbb{V}_k[\varphi] \right| < \infty,$$

it follows that the remainder of the integral in (84) can be bounded by $\varepsilon(1 - e^{-\lambda(k-1)t})/\lambda(k-1)$, which can be bounded by ε . Combining this with (82) we get the desired result. \square

Proof of Theorem 3 (subcritical case). We now outline the proof for the subcritical case. Again we use an inductive argument. The case $k = 1$ follows from (H1) and the fact that $\langle \varphi, \tilde{\varphi} \rangle = 1$. Now assume the result to be true for $\ell = 1, \dots, k-1$. We first note first that the term $R_k(x, t)$ in (69) vanishes in the limit after the normalisation $e^{-\lambda t}$. To see this, note from (65) that

$$\begin{aligned}
\left| \frac{e^{-\lambda t}}{\varphi(x)} R_k(x, t) \right| &\leq \sum_{\{m_1, \dots, m_{k-1}\}_k} c(m_1, \dots, m_{k-1}) \prod_{j=1}^{k-1} \left| \frac{e^{-\lambda t} \mathbf{T}_t^{(j)}[f](x)}{\varphi(x) j!} \right|^{m_j} \\
&\quad \times \varphi(x)^{m_1 + \dots + m_{k-1}} e^{\lambda(m_1 + \dots + m_{k-1} - 1)t},
\end{aligned}$$

where $c(m_1, \dots, m_{k-1})$ is a constant depending only on m_1, \dots, m_{k-1} . Since each of the terms in the product is bounded, $\lambda < 0$, and $m_1 + \dots + m_{k-1} > 1$ for any partition, the limit of the right-hand side above is zero. Using this, along with the induction hypothesis, assumption (H1) and the evolution equation (69) we see that

$$\begin{aligned}
(85) \quad &\lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \varphi(x)^{-1} e^{-\lambda t} \mathbf{T}_t^{(k)}[f](x) - k! \langle f, \tilde{\varphi} \rangle^k \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right| \\
&= \lim_{t \rightarrow \infty} \sup_{x \in E, f \in B_1^+(E)} \left| \frac{e^{-\lambda t}}{\varphi(x)} \int_0^t \mathbf{T}_s[U_k^*(\cdot, t-s)](x) ds - k! \langle f, \tilde{\varphi} \rangle^k \langle \mathbb{V}_k[\varphi], \tilde{\varphi} \rangle \right|,
\end{aligned}$$

where

$$U_k^*(x, t) = \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \left[\psi^{(m_1 + \dots + m_{k-1})}(x, 0+) \prod_{j=1}^{k-1} \left(-\frac{\mathbf{T}_t^{(j)}[f](x)}{j!} \right)^{m_j} + \beta(x) \int_{M(E)^\circ} \prod_{j=1}^{k-1} \left(\frac{1}{j!} \langle \mathbf{T}_t^{(j)}[f], \nu \rangle \right)^{m_j} \Gamma(x, d\nu) \right].$$

Then, similar calculations to those above yield

$$\begin{aligned} \frac{e^{-\lambda t}}{\varphi(x)} \int_0^t \mathbf{T}_s[U_k^*(\cdot, t-s)](x) ds &= t \sum_{\{m_1, \dots, m_{k-1}\}_k} \frac{k!}{m_1! \dots m_{k-1}!} \int_0^1 e^{-\lambda t(1-u)(1-m_1-\dots-m_{k-1})} \\ &\frac{e^{-\lambda tu}}{\varphi(x)} \mathbf{T}_{tu} \left[\psi^{(m_1 + \dots + m_{k-1})}(\cdot, 0+) (-\varphi(\cdot))^{m_1 + \dots + m_{k-1}} \prod_{j=1}^{k-1} \left(\frac{e^{-\lambda t(1-u)} \mathbf{T}_{t(1-u)}^{(j)}[f](\cdot)}{\varphi(\cdot) j!} \right)^{m_j} \right. \\ &\left. + \beta(\cdot) \int_{M(E)^\circ} \prod_{j=1}^{k-1} \left\langle \varphi \frac{e^{-\lambda t(1-u)}}{\varphi j!} \mathbf{T}_t^{(j)}[f], \nu \right\rangle^{m_j} \Gamma(x, d\nu) \right] (x) du. \end{aligned}$$

To finish the proof, we use induction hypothesis to deal with the lower order moments, dominated convergence and (H1) to deal with the limit of $\mathbf{T}_{t(1-u)}$. Similarly to the last part of the proof of Theorem 2 we get that (85) is bounded as $t \rightarrow \infty$ by $\varepsilon(1 - e^{-\lambda t(1-m_1-\dots-m_k)})/\lambda(1-m_1-\dots-m_k)$, which is bounded by ε since $m_1 + \dots + m_{k-1} > 1$. We once more leave the details of the rest of the proof to the reader. \square

Proofs of Theorems 4, 5 and 6. Given the proofs we have now seen for the branching particle setting for these three theorems, as well as the proofs of Theorem 1, 2 and 3, we mention only that a similar calculation to the one presented in Theorem 7 tells us that

$$\begin{aligned} \mathbf{M}_t^{(k)}[g](x) &= (-1)^{k+1} \tilde{R}_k(x, t) + (-1)^k \int_0^t \mathbf{T}_s \left[\tilde{U}_k(\cdot, t-s) \right] (x) ds \\ &\quad - k \int_0^t \mathbf{T}_s \left[g[\mathbf{M}_{t-s}^{(k-1)}[g]] + (-1)^{k-1} \tilde{R}_{k-1}(\cdot, t-s) \right] (x) ds, \end{aligned}$$

where

$$\tilde{U}_k(x, t) = \tilde{K}_k(x, t) + \beta(x) \tilde{S}_k(x, t)$$

and \tilde{R} , \tilde{K} and \tilde{S} are defined as R , K and S in (65), (66), (67), respectively, albeit replacing $\mathbf{T}^{(j)}$ by $\mathbf{M}^{(j)}$. From here we can consider the claimed asymptotics using the inductive reasoning as in the proof of Theorems 1, 2 and 3, respectively. \square

Appendix

We first state two fundamental combinatorial results for complex derivatives, the classical Faà di Bruno and Leibniz rules. In both cases, for a sufficiently smooth function g on \mathbb{R} , we will denote by $g^{(k)}$ by its k -th derivative.

Lemma A.1 (Faà di Bruno rule). *Let f and g k -times continuously differentiable functions on \mathbb{R} . Then the k -th derivative is given by the following formula*

$$\frac{d^k}{dx^k} f(g(x)) = \sum_{\{m_1, \dots, m_k\}_k} \frac{k!}{m_1! \dots m_k!} f^{(m_1 + \dots + m_k)}(g(x)) \prod_{j=1}^k \left(\frac{g^{(j)}(x)}{j!} \right)^{m_j},$$

where the sum goes over the set $\{m_1, \dots, m_k\}_k$ of non-negative integers such that

$$m_1 + 2m_2 + \dots + km_k = k.$$

Lemma A.2 (Leibniz rule). *Suppose g_1, \dots, g_m are k -times continuously differentiable functions on \mathbb{R} , for $k \geq 1$. Then*

$$\frac{d^k}{dx^k} \left(\prod_{i=1}^m g_i(x) \right) = \sum_{k_1 + \dots + k_m = k} \binom{k}{k_1, \dots, k_m} \prod_{\ell=1}^m g_\ell^{(k_\ell)}(x),$$

where the sum is taken over all non-negative integers k_1, \dots, k_m such that $\sum_{i=1}^m k_i = k$.

The third result of the appendix is a general ergodic limit theorem which is key to the moment convergence. We will only state the result in the critical case, since we will only apply it in the proof of Theorem 1, however, the result can easily be extended to the non-critical case by including the normalisation $e^{-\lambda ut}$ in the first integral.

In order to state it, let us introduce a class of functions \mathcal{C} on $B_1^+(E) \times E \times [0, 1] \times [0, \infty)$ such that F belongs to class \mathcal{C} if

$$F[g](x, s) := \lim_{t \rightarrow \infty} F[g](x, s, t), \quad g \in B_1^+(E), x \in E, s \in [0, 1],$$

exists,

$$(A.1) \quad \sup_{x \in E, s \in [0, 1], g \in B_1^+(E)} |\varphi F[g](x, s)| < \infty,$$

and

$$(A.2) \quad \lim_{t \rightarrow \infty} \sup_{x \in E, s \in [0, 1], g \in B_1^+(E)} \varphi(x) |F[g](x, s) - F[g](x, s, t)| = 0.$$

Theorem A.1. *Assume (H1) holds, $\lambda = 0$ and that $F \in \mathcal{C}$. Define*

$$\Xi_t = \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \int_0^1 \mathbb{T}_{ut}[\varphi F[g](\cdot, u, t)](x) du - \int_0^1 \langle F[g](\cdot, u), \varphi \tilde{\varphi} \rangle du \right|, \quad t \geq 0.$$

Then

$$(A.3) \quad \sup_{t \geq 0} \Xi_t < \infty \text{ and } \lim_{t \rightarrow \infty} \Xi_t = 0.$$

Proof. We will show that

$$\lim_{t \rightarrow \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \mathbf{T}_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi} \varphi, F[g](\cdot, u) \rangle \right| = 0,$$

since then

$$\begin{aligned} \lim_{t \rightarrow \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \int_0^1 \frac{1}{\varphi(x)} \mathbf{T}_{ut}[\varphi F[g](\cdot, u, t)](x) du - \int_0^1 \langle \tilde{\varphi} \varphi, F[g](\cdot, u) \rangle du \right| \\ \leq \int_0^1 \lim_{t \rightarrow \infty} \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \mathbf{T}_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi} \varphi, F[g](\cdot, u) \rangle \right| du = 0. \end{aligned}$$

First note that,

$$\begin{aligned} & \left| \frac{1}{\varphi(x)} \mathbf{T}_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi} \varphi, F[g](\cdot, u) \rangle \right| \\ & \leq \frac{1}{\varphi(x)} \mathbf{T}_{ut}[|\varphi F[g](\cdot, u, t) - \varphi F[g](\cdot, u)|](x) \\ & \quad + \left| \frac{1}{\varphi(x)} \mathbf{T}_{ut}[\varphi F[g](\cdot, u)](x) - \langle \tilde{\varphi}, \varphi F[g](\cdot, u) \rangle \right| \end{aligned}$$

Due to assumption (A.2), for t sufficiently large, the first term on the right-hand side above can be controlled by $\varphi^{-1}(x) \mathbf{T}_{ut}[\varepsilon](x)$. Combining this with the above inequality yields

$$\begin{aligned} \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \mathbf{T}_{ut}[\varphi F[g](\cdot, u, t)](x) - \langle \tilde{\varphi} \varphi, F[g](\cdot, u) \rangle \right| \\ \leq \sup_{x \in E} |\varphi^{-1}(x) \mathbf{T}_{ut}[\varepsilon](x) - \langle \tilde{\varphi}, \varepsilon \rangle| + \varepsilon \|\tilde{\varphi}\|_1 \\ \text{(A.4)} \quad + \sup_{x \in E, g \in B_1^+(E)} \left| \frac{1}{\varphi(x)} \mathbf{T}_{ut}[\varphi F[g](\cdot, u)](x) - \langle \tilde{\varphi}, \varphi F[g](\cdot, u) \rangle \right|. \end{aligned}$$

We note that (A.1) and the first (resp. second) statement of (13) in (H1), together with dominated convergence, immediately imply that the first (resp. second) statement in (A.3) holds. \square

References

- [1] Z-Q. Chen, Y-X. Ren, and T. Yang. Law of large numbers for branching symmetric Hunt processes with measure-valued branching rates. *J. Theoret. Probab.*, 30(3):898–931, 2017.
- [2] Zhen-Qing Chen, Yan-Xia Ren, and Ting Yang. Skeleton decomposition and law of large numbers for supercritical superprocesses. *Acta Appl. Math.*, 159:225–285, 2019.

- [3] D. A. Dawson. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of *Lecture Notes in Math.*, pages 1–260. Springer, Berlin, 1993.
- [4] Donald A Dawson, Luis G Gorostiza, and Zenghu Li. Nonlocal branching superprocesses and some related models. *Acta Applicandae Mathematica*, 74(1):93–112, 2002.
- [5] E. Dumonteil and A. Mazzolo. Residence times of branching diffusion processes. *Phys. Rev. E*, 94:012131, 2016.
- [6] E. Dynkin. *Diffusions, superdiffusions, and partial differential equations*, volume 50. American Mathematical Society Providence, RI, 2002.
- [7] E. B. Dynkin. Branching particle systems and superprocesses. *Ann. Probab.*, 19(3):1157–1194, 1991.
- [8] Maren Eckhoff, Andreas E. Kyprianou, and Matthias Winkel. Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Probab.*, 43(5):2545–2610, 2015.
- [9] J. Engländer. *Spatial branching in random environments and with interaction*, volume 20 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015.
- [10] J. Engländer, S. C. Harris, and A. E. Kyprianou. Strong law of large numbers for branching diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):279–298, 2010.
- [11] A. M. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.
- [12] J. Fleischman. Limiting distributions for branching random fields. *Trans. Amer. Math. Soc.*, 239:353–389, 1978.
- [13] C. S. Harris, E. Horton, A. E. Kyprianou, and M. Wang. Yaglom limit for critical neutron transport, 2021.
- [14] S. C. Harris, E. Horton, and A. E. Kyprianou. Stochastic methods for the neutron transport equation II: almost sure growth. *Ann. Appl. Probab.*, 30(6):2815–2845, 2020.
- [15] S. C. Harris and M. I. Roberts. The many-to-few lemma and multiple spines. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(1):226–242, 2017.
- [16] E. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic methods for the neutron transport equation I: linear semigroup asymptotics. *Ann. Appl. Probab.*, 30(6):2573–2612, 2020.
- [17] I. Iscoe. On the supports of measure-valued critical branching Brownian motion. *Ann. Probab.*, 16(1):200–221, 1988.
- [18] A. Klenke. Multiple scale analysis of clusters in spatial branching models. *Ann. Probab.*, 25(4):1670–1711, 1997.

- [19] Z. Li. *Measure Valued Branching Markov Processes*. Springer-Verlag Berlin Heidelberg, 2011.
- [20] S. Palau and T. Yang. Law of large numbers for supercritical superprocesses with non-local branching. *Stochastic Process. Appl.*, 130(2):1074–1102, 2020.
- [21] Y-X. Ren, R. Song, and Z. Sun. Limit theorems for a class of critical superprocesses with stable branching. *Stochastic Process. Appl.*, 130(7):4358–4391, 2020.
- [22] B. A. Sevast'janov. The extinction conditions for branching processes with diffusion. *Teor. Veroyatnost. i Primenen.*, 6:276–286, 1961.
- [23] B. A. Sevast'yanov. Branching stochastic processes for particles diffusing in a bounded domain with absorbing boundaries. *Teor. Veroyatnost. i Primenen.*, 3:121–136, 1958.
- [24] A. V. Skorohod. Branching diffusion processes. *Teor. Veroyatnost. i Primenen.*, 9:492–497, 1964.
- [25] S. Watanabe. A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.*, 8:141–167, 1968.